In re Application of:)	
)	Group Art Unit: 1645
ZHELNIN and BLOOMQUIST	Ś	•
	Ś	Examiner: S. Gucker
Serial No. 09/899,532	Ś	
50Hai 140. 07/077,332	, ,	
THE T . T . C . C . C . C . C . C . C . C .	,	D 1 - 31 - 00070 00040
Filing Date: July 6, 2001)	Docket No. 02973.00040

For: HUMAN NEUROPEPTIDE Y-LIKE G PROTEIN-COUPLED RECEPTOR

DECLARATION UNDER 37 C.F.R. § 1.131

TED STATES PATENT AND TRADEMARK OFFICE

U.S. Patent and Trademark Office 220 20th Street S. Customer Window, Mail Stop AF Crystal Plaza Two, Lobby, Room 1B03 Arlington, VA 22202

Dear Sir:

We, Leonid Zhelnin and Brian Bloomquist, declare as follows:

- 1. We are the named inventors of the subject matter claimed in the application referenced above.
- 2. Prior to March 2, 2000, we reduced to practice the subject matter of claims 1-8 and 12 of the application referenced above. The dates on the attached exhibits have been reducted; however, all the work described in this declaration was performed in the United States prior to March 2, 2000.
- 3. Prior to March 2, 2000, we searched for sequences that could be used to clone novel GPCRs with homology to orexin-2 and NPY1 receptors. We identified these sequences by

monitoring search results from a program (Gene Agent) that automatically checks sequences released into the public domain for homology to the orexin-2 and NPY1 receptors.

- 4. Prior to March 2, 2000, we received reports from both the Gene Agent program orexin-2 and NPY1 receptor searches indicating that DNAs within accession numbers EM:AC005961.1 (Homo sapiens Chromosome 11q12.2 PAC pDJ32L16 genomic DNA) and GB:g3823006 (Human sperm genomic library; Accession number AQ270411) each had homology to orexin-2 and NPY1 receptor sequences. Copies of the reports are attached as Exhibits 1 and 2.
- 5. Exhibit 1 identifies G3823006 as having homology to the orexin-2 receptor sequence. Pages 1 through 3 of Exhibit 1 contain analysis results from the search for sequences with homology to the orexin-2 receptor. See the top of page 1, which describes the "Project Name" as "orexin receptor homology search." In the "Protein Agent" box near the bottom of page 1, the location of G3823006 ("GB:g3823006") relative to the orexin-2 receptor sequence (OX2R) is shown. Pages 2 and 3 of Exhibit 1 contain initial information about G3823006 (beginning with the last set of boxes on page 2). Pages 4 and 5 of Exhibit 1 provide more detailed information on G3823006, including its length of 483 nucleotides.
- 6. Exhibit 2 identifies AC005961 as having homology to the orexin-2 receptor sequence. Pages 1 and 2 of Exhibit 2 also contain analysis results from the search for sequences with homology to the orexin-2 receptor. See the top of page 1, which describes the "Project Name" as "orexin receptor homology search." Page 2 of Exhibit 2 specifically contains information about AC005961 (last set of boxes on page 2). Page 3 of Exhibit 2 contains a summary of the search results for sequences with homology to the orexin-2 receptor. See the top of page 3, which indicates that the "Project Name" is "orexin receptor homology search." The

middle portion of page 3 ("Sequences") contains a short description of each of the seven sequences identified by the search. These sequences include both accession numbers EM:AC005961 and GB:g3823006. Page 4 of Exhibit 1 provides additional information about accession number AC005961.1, including its length of 154,647 nucleotides.

- 7. We analyzed the sequences of these DNAs and found them to be overlapping. We selected the DNA sequence of Accession Number AC005961 to continue our efforts to obtain the sequence of our novel orexin- or NPY-like receptor (hereafter referred to as the "novel GPCR").
- 8. BLAST analysis revealed that within AC005961 there were sequences comprising at least five exons with homology to the orexin-2 receptor. Exhibit 3 provides a copy of BLAST comparisons of five potential exons of an amino acid sequence of a putative novel GPCR encoded within Accession Number AC005961 (Sbjct) relative to the human orexin-2 receptor (Query) at page 1. Pages 2-3 of Exhibit 3 contain the description of the human orexin-2 receptor query sequence as found under Accession Number NP_001517.
- 9. Prior to March 2, 2000, we used sequence information from AC005961 to begin cloning the novel GPCR. We designed primers to amplify a portion of the novel GPCR from human genomic DNA using the polymerase chain reaction (PCR). We obtained PCR products of 161 bp ("NSAN.seq" or "GA1") and 234 bp ("GA2") from amplification of the genomic DNA. The 161 bp amplicon included DNA from transmembrane domains (TM) 1 and 2 of the putative novel GPCR. The 234 bp product included sequence from TM 3 and 4 domains of the putative novel GPCR.
- 10. A copy of the nucleotide sequences of the two PCR products is attached as Exhibit 4. Page 1 of Exhibit 4 describes the 161 bp sequence, which is identified by the

Sequence ID "NSAN.seq", as the "pcr of g3823006 GA1." Page 2 of Exhibit 4 shows the sequence of "NSAN.seq." Nucleotides 11-153 of the 161 base pair "NSAN.seq" amplification product are identical to nucleotides 167-309 of claimed SEQ ID NO: 1. Exhibit 5 provides an alignment of the "NSAN.seq" sequence with SEQ ID NO: 1. The first 10 and last 8 nucleotides of the 161 bp clone, which do not align with SEQ ID NO: 1, are not exonic sequence.

- 11. A description of the 234 bp amplification product, identified as "GA2" (see the Sequence ID), is provided on page 3 of Exhibit 4. Page 4 of Exhibit 4 provides the sequence of "GA2." Nucleotides 29-191 of the sequence of the "GA2" are identical to nucleotides 340-502 of claimed SEQ ID NO: 1. See Exhibit 6 for an alignment of "GA2 with SEQ ID NO: 1. The first 28 and last 43 nucleotides of the 234 bp clone do not align and presumably represent intronic sequence.
- 12. We designed primers to the 5'- and 3'-regions of the 161 bp amplification product, and performed 5'- and 3'-RACE to begin our effort to obtain the full-length nucleotide sequence of the novel GPCR. The reverse primers used in the strategy to extend the 5'-sequence of the novel GPCR were termed "GA1-85B" and "GA1-126B." These primers are shown on pages 2 and 3 of Leonid Zhelnin's laboratory notebook no. RB53651; copies of Leonid Zhelnin's laboratory notebook pages 2 and 3 are provided as Exhibit 7.
- 13. The sequence of the "GA1-85B" primer is shown on the right side at the top of page 2 of Leonid Zhelnin's laboratory notebook. The sequence of the "GA1-126B" primer is shown on the right side at the top of page 2, beneath the sequence of the "GA1-85B" primer. Page 2 also shows the location of primers "GA1-85B" and "GA1-126B" on the 161 nucleotide sequence of "NSAN.seq". The sequence of each primer on the 161 nucleotide sequence is

circled, and "GA1-85B" and "GA1-126B" is handwritten above the location of each primer on the sequence.

- 14. Page 3 of Leonid Zhelnin's laboratory notebook shows the location of the primers "GA1-85B" and "GA1-126B" on an alignment of the 161 bp subregion of G3823006, "NSAN.seq" ("GA1"). The sequence of each primer on the 161 nucleotide sequence is circled, and "GA1-85B" and "GA1-126B" is handwritten below the location of each primer on the sequence. The locations of nested PCR primers are also indicated.
- 15. Exhibit 8 shows the location of the "GA1-85B" and "GA1-126B" RACE primers against SEQ ID NO:1. The shaded box indicates the position of the 161 bp fragment, NSAN.seq; the identities of amino acids of the codons represented within the primers are included for reference purposes.
- 16. Prior to March 2, 2000, using the "GA1-85B" and "GA1-126B" primers, we obtained two 5'-RACE clones (210 and 347 bp) from human heart cDNA that were used to extend the sequence further 5' to NSAN.seq (GA1). Assembly of sequence data from the 5'-RACE clones and the 161 bp clone, NSAN.seq, revealed a 347 bp long ORF, named 5GA1. The predicted 103 amino-acid peptide encoded by this ORF had significant homology to the N-terminal regions of the NPY-1 and orexin-2 receptors.
- 17. The sequence of the nucleotide and encoded polypeptide of the 347 bp ORF, 5GA1, are provided in Exhibit 9, which contains a copy of pages 47 and 48 of Leonid Zhelnin's laboratory notebook no. RB53651. Page 47 (page 1 of Exhibit 9), shows the nucleotide and encoded polypeptide sequence of the 347 bp RACE clone (top), and a ClustalW-formatted alignment of the predicted amino acid sequence of 5GA1 with the amino acid sequence of the neuropeptide Y1 receptor (NPY-1R.pep) (bottom). Page 48 (page 2 of Exhibit 9) shows a

ClustalW-formatted alignment of the predicted amino acid sequence of 5GA1 with the amino acid sequence of the orexin-2 receptor (Orexin 2.pep).

- 18. An alignment of the amino acid sequence encoded by 5GA1 with SEQ ID NO: 2 is provided in Exhibit 10. The alignment indicates that the amino acid sequence encoded by 5GA1 is identical to amino acid residues 1-103 of SEQ ID NO: 2. The last two predicted residues of 5GA1 are not represented in SEQ ID NO: 2.
- 19. An alignment of the nucleotide sequence of 5GA1 to SEQ ID NO: 1 is provided in Exhibit 11. The alignment shows that the nucleotide sequence of the 5GA1 clone is identical to nucleotides 1 309 of SEQ ID NO: 1. The nucleotide sequence of 5GA1 further extended the 161 bp sequence of NSAN.seq, the initial PCR fragment identical to nucleotides 167-309 of SEQ ID NO: 1 (see paragraph 6 and alignment in Exhibit 5), to include nucleotides 1 166 of SEQ ID NO: 1.
- 20. The position of the putative START methionine codon (ATG) in clone 5GA1 was consistent with clone 5GA1 being similar to the NPY-1 and orexin-2 receptors. The identity of this ATG as the true START codon was, however, unconfirmed at this point in the cloning because the open reading frame (ORF) was not closed upstream of the 5'-most ATG.
- Documentation of the initial efforts to perform 3'-RACE are provided as Exhibit 12, which contains copies of pages 4 and 5 of Leonid Zhelnin's laboratory notebook no. RB53651. These pages show the location of the "forward" primers termed "GA1-F32" and "GA1-F41." The nucleotide sequence of primer "GA1-F32" is shown at the bottom of page 4 of Leonid Zhelnin's laboratory notebook. The nucleotide sequence of primer "GA1-F41" is shown as the nucleotide sequence immediately above the nucleotide sequence of primer "GA1-F32."

- 22. Page 4 of Leonid Zhelnin's laboratory notebook, the first page of Exhibit 12, also shows the location of each of primers "GA1-F32" and "GA1-F41" in the sequence of the 161 nucleotide clone ("NSAN.seq" or "GA1"). The nucleotide sequence of each primer is circled on "NSAN.seq." "GA1F41" is also handwritten above its nucleotide sequence in "NSAN.seq." "6AF32" is handwritten to the side of the nucleotide sequence for "GA1-F32". Exhibit 8, described above in paragraph 15, also diagrams the location of the forward primers relative to SEO ID NO: 1.
- 23. The amplification conditions for performing 3'-RACE are provided on page 5 of Leonid Zhelnin's laboratory notebook, which is page 2 of Exhibit 12. The product of 3'-RACE can be seen in a photograph of an agarose gel provided at the bottom right-hand corner of page 1 of Exhibit 12, see particularly lanes 3 and 4 from the left.
- 24. We analyzed the sequence of the 3'-RACE amplicon procedure, which revealed that this clone contained sequences homologous to a GPCR through the seventh transmembrane domain, and ended in a polyadenosine stretch but with no STOP codon. Exhibit 13 is an alignment of two independent 3'-RACE amplicons, LZLZ11 and LZLZ21, with the "5GA1" consensus sequence as recorded on page 53 of Leonid Zhelnin's laboratory notebook no. RB53651.
- 25. LZLZ11 and LZLZ21 are 650 bp and 730 bp in length, respectively. When these two amplicons of overlapping sequence (see Exhibit 13, described in paragraph 24 above) are linked through PCR amplification they yielded a clone (3'-RACE #2). The 3'-RACE #2 clone is 913 bp in length. An alignment of 3'-RACE #2 clone with SEQ ID NO: 1 is shown in Exhibit 14A: The 913 base pair sequence of the 3'-RACE #2 clone is identical to nucleotides 230-1017 of SEQ ID NO: 1. An alignment of the amino acid sequence encoded by 3'-RACE #2 with SEQ

- ID NO: 2 is shown in Exhibit 14B. The amino acid sequence encoded by 3'-RACE #2 is identical to amino acid residues 78-339 of SEQ ID NO: 2.
- 26. Assembling, by computer, the sequence of 3'-RACE #2 clone and clone 5GA1, we predicated a hypothetical 1036-nucleotide sequence encoding 345 amino acids. The sequence of the "virtual" 1036-nucleotide clone is identical to nucleotides 1-1017 of SEQ ID NO: 1. An alignment of these sequences is shown in Exhibit 15. The nucleotide sequences of the individual LZLZ11, LZLZ21 and 5GA1 sequences, which comprise the 3'-RACE #2 clone, are separately highlighted in Exhibit 15, pages 1 and 2.
- The apparent poly(A) tail of the 3'-RACE #2 clone is not found in SEQ ID NO: 1. The 345 amino acid sequence encoded by the virtual clone is identical to amino acid residues 1-339 of SEQ ID NO: 2. An alignment of the 345-amino acid sequence with SEQ ID NO: 2 is provided on page 3 of Exhibit 15. Given the fact that the 3'-RACE #2 clone (and, hence, the virtual clone) had no STOP codon and an abnormally short C-terminal intracellular domain, we predicted that we still needed to identify more 3' sequence to complete the clone. The potential STOP codon highlighted at the bottom right of page 53 of Leonid Zhelnin's notebook no. RB53651 (Exhibit 13) actually is part of the Marathon-Ready™ cDNA adaptor in the vector.
- 28. Prior to March 2, 2000, we used the virtual sequence shown in Exhibit 13 as a template to search the Image Consortium database for novel expressed sequence tags (ESTs) to identify the missing 3'-nucleotide sequence. We identified three Image clones (named Im: 2055532, Im: 2055185, and Im: 2091776) from human kidney which had similarity to the virtual clone. Identification of the three Image clones was recorded at the top of page 61 of Leonid Zhelnin's laboratory notebook no. RB53651 (Exhibit 16).

- 29. Sequence analysis of Image clone 2055185 confirmed the putative START methionine described in paragraph 20 because it revealed an in-frame STOP codon upstream of this START methionine. Page 89 of Leonid Zhelnin's laboratory notebook no. RB53651 shows the START an STOP codons of the assembled sequence (Exhibit 17). The in-frame "TAG" STOP codon is boxed and begins 111 bp upstream of the first nucleotide of the START methionine, which is marked with an "M" at nucleotide 394 of the sequence.
- 30. We again performed a BLAST search of the sequence contained in AC005961 (as described in paragraph 4 above) to obtain the 3'-most sequence of the novel GPCR. We used a short stretch of sequence as query (about twenty nucleotides). The query sequence was located between the site of a several hundred-nucleotide unspliced intron and a stretch of 8 adenosine residues found in the sequence near the 3'-end of one of the Image clones. The top portion of Exhibit 18 diagrams the BLAST search strategy. Done this way, it was hoped that the search algorithm would not miss critical homologies with this short region which might otherwise be lost because of relevant sequence being "diluted" within a larger query sequence potentially broken up by other unidentified unspliced introns.
- 31. BLAST analysis identified the 279 nucleotide region that contained the putatively complete 3'-end and STOP codon of the novel GPCR. This information is provided as nucleotides 1411-1689 of the sequence on page 89 of Leonid Zhelnin's laboratory notebook no. RB53651 (Exhibit 19 is a copy of page 89 of Leonid Zhelnin's laboratory notebook, the full-length sequence is referred to as LGZ1.seq).
- 32. The 279 nucleotides are identical to nucleotides 1018-1296 of SEQ ID NO: 1. An alignment of the 3'-most 300 nucleotides (279 bp of coding and 21 bp of 3'-untranslated) of the full-length novel GPCR, LGZ1, with SEQ ID NO:1 is provided in Exhibit 20.

- 33. The coding region of the novel GPCR included 1,296 nucleotides separated into 6 exons within the 154,647 nucleotides of genomic sequence AC005961. Exhibit 18, bottom diagram, shows the genomic structure of the novel GPCR.
- 34. Prior to March 2, 2000, we amplified clones containing the complete 1296-bp ORF as well as 5'- and 3'-untranslated DNA from both human heart and brain cDNA by PCR (using Marathon-Ready cDNA libraries, Clontech). The complete sequence of the novel GPCR was recorded on page 89 of Leonid Zhelnin's laboratory notebook no. RB53651. Exhibit 17, also described in paragraph 29, provides a copy of page 89 of Leonid Zhelnin's laboratory notebook. The 1296-bp ORF is identical to SEQ ID NO:1.
- 35. Also prior to March 2, 2000, we confirmed the coding sequence of the full-length cDNA clone by sequence analysis of a full-length amplicon generated by PCR from human brain cDNA. This sequence was recorded on pages 92 and 96-97 of Leonid Zhelnin's laboratory notebook no. RB53651; copies of these pages are provided as Exhibit 21.
- 36. The coding sequence obtained by the sequence analysis is contained within SEQ ID NO:1 of our provisional application Serial No. 60/216,523, which is identical to SEQ ID NO:1 of our utility application 09/899,532. Exhibit 22 shows the 100% identity between SEQ ID NO: 1 and nucleotides 19 1335 of the sequence shown in Exhibit 21.
- 37. The amino acid sequence encoded by nucleotides 19 1335 in Exhibit 21 is identical to SEQ ID NO: 2 of our provisional application Serial No. 60/216,523, which is identical to SEQ ID NO: 2 of our utility application 09/899,523. An alignment of SEQ ID NO: 2 of the provisional and utility applications with the amino acid sequence shown in Exhibit 21 is provided in Exhibit 23. In addition, an alignment of the amino acid sequence shown in Exhibit 21 showing the constructs used to assemble SEQ ID NO: 2 is shown in Exhibit 24.

All statements made herein of our own knowledge are true and all statements 38. made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Dated: January 31, 2005

Brian T. Bloomquist, Ph.D.

Dated: January 31,2005 Leonid Zhelnin

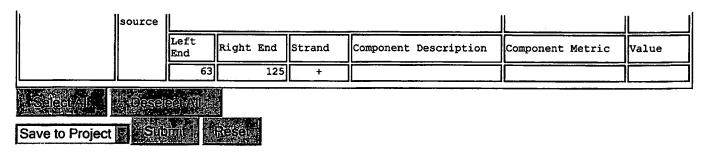
Exhibit 1 Identification of G3823006



Analysis Results

, eres	Main Me	nu nu	Projects	ac maj	jue ry	Seq Analysi	a. Utilities	The same of the sa	login / Log	out.	Help			
Cr Mo St	oject I eated: dified: atus: day's D	OP	exin En	C M C	reated By odified By	: Zhelnin, y: lt_Accou			search					
	nalysis Da		ved. Use the	buttons to	o change the	state of these	results.							
De	escription		ar Pogus		A Direct	io presenta	sults							
Tr Fe	ans. Prote	in: 5 -up: C	Off © On		55		·							
						Prote	in Agent							
					mR	con ene INA ource			- - GB : g382		8 : g471120 : GB : g3820)	0X2R	
	L				GB:g38230		xin:OX2R:1							
	Hit	ID	P-Value	% ident.			Descri	ption	1				Hit	Class
ľ	V		8.8e-19	34	Human mRN	A for angle	otensin II ty	pe 11	recepto	or, co	mplete c	ds.	Tran Prot	
	GB:g471 Result D	120:2 etails	Feature			Descrip	tion			Fea	ture Met	ric	Va	lue

	CDS	L					
		Left End	Right End	Strand	Component Description	Component Metric	Value
		61	160	+			
	source						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		61	160	+]
Hit ID	P-Value	% ident.			Description		Hit Class
☑ GB:g3820434:2	2.2e-18	30	Mus musculu CDS.	is mrna fo	r G-protein coupled rece	eptor, complete	Trans. Protein
Result Details	Feature			Descrip	tion	Feature Metric	Value
	CDS		ir	1			
		Left End	Right End	Strand	Component Description	Component Metric	Value
		62	162	+][][
	gene						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		62	162	+			
	source						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		62	162	+			
Hit ID	P-Value	% ident.		Description			
☑ GB:q3822556:1	3.6e-14	33	Bos taurus cds.	oxytocin :	receptor gene, promoter	region and partial	Trans. Protein
Result Details				Descrip	tion	Feature Metric	Value
	CDS	7 - 61		·			
		Left End	Right End	Strand	Component Description	Component Metric	Value
		56	160	+			
	exon						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		56	160	+			
	mRNA						
*		Left End	Right End	Strand	Component Description	Component Metric	Value
		56	160	+			
	source						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		56	160	+			
Hit ID	P-Value	% ident.			Description		Hit Class
☑ GB:g3823006:1	9.6e-10	39	Homo sapier	ns genomic	T Approved Human Genomic clone Plate=2052 Col=7		Trans. Protein
Result Details	Feature		survey sequ	Descrip	tion	Feature Metric	Value
	COMPLETION			LescIID:	LAVII	" Beerme Merric	



Requested by: Zhelnin, Leonid



INCYTE PHARMACEUTICALS, INC.



Query Detail

Utilities Main Menu Projects Seq Analysis Login / Logout Query Help

Sequence: GB:g3823006:1



Entries with the same Physical Sequence Newgage:g3823006:1 orexin:g3823006:1									
General Information About The Entry									
Sequence ID	g3823006								
Version	1								
Sequence Description	HS_2052_B1_G04_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2052 Col=7 Row=N, genomic survey sequence.								
Database	GB GENBANK								
Database Description	Nucleotide and Protein Sequence Data Bank								
Molecule	DNA								
Database Division	GSS								
Entry Header Line	AQ270411 483 bp DNA GSS								
Alternate Keys	LOCUS AQ270411 PRIMARY AC AQ270411								
Sequence Inform	ation								
Length	483								
Sequence Header Line	155 a 84 c 169 g 73 t 2 others								
Origin and Alte	rnate Names								
Keywords	GSS								
Organism	Homo sapiens								
Taxonomy	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
Species	HOMO SAPIENS								
Biological Tissue Source	human.								
Comments									

- Contact: Mahairas GG, Wallace JC, Hood L
- · High Throughput Sequencing Center
- University of Washington
- 401 Queen Anne Avenue North, Seattle, WA 98109, USA
- Tel: (206) 616-3618 Fax: (206) 616-3887
- Email: jwallace@u.washington.edu
- Sequence Tagged Connector
- Plate: 2052 row: N column: 7
- Class: BAC ends

(bases 1 to 483) Construction of a Characterized Clone Resource for Genomic Sequencing Unpublished Mahairas G.G. Wallace J.C. Smith K. Swartzell S. Furlong J. Shaker R. Schmidt S. Traicoff R. Hood L.E.
Unpublished Mahairas G.G. Wallace J.C. Smith K. Swartzell S. Furlong J. Shaker R. Schmidt S. Traicoff R. Hood L.E.
Mahairas G.G. Wallace J.C. Smith K. Swartzell S. Furlong J. Shaker R. Schmidt S. Traicoff R. Hood L.E.
Wallace J.C. Smith K. Swartzell S. Furlong J. Shaker R. Schmidt S. Traicoff R. Hood L.E.
ires
Gualifiers Plate=2052 Col=7 Row=N ib CIT Approved Human Genomic Sperm Library D Homo sapiens male Cross References 9606 Notes sperm; Vector: pBeloBAC11; BAC Clones in
i



INCYTE PHARMACEUTICALS, INC.

Exhibit 2 Identification of AC005961



Analysis Results

Main Menu Projects Overy Seq Analysis Utilities Login / Logout Help
Project ID: orexin Project Name:orexin receptor homology search Created: Created By: Zhelnin, Leonid Modified: Modified By: lt_Account, Admin Status: OPEN Current User:Zhelnin, Leonid (WRITE) Today's Date: Num Seqs: 7
Analysis Date:
These results are saved. Use the buttons to change the state of these results.
Save These results 2 Apriscard these Casults 3 Description:
Maximum Number of Hits per Hit Class Trans. Protein: 5 Feature roll-up: C Off © On Hedispla
View parameters
Protein Agent
□ orexin:OX2R:1

Hit ID	P-Value	% ident.	Description				
	2.5e-17		Caenorhabdi	aenorhabditis elegans cosmid T05Al			
Result Details	Feature			Descri	otion	Feature Metric	Value
	source						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		81	134	+			
Hit ID	P-Value	% ident.		Description			
EM: CEW05B5:5	4.7e-17			aenorhabditis elegans cosmid W05B5			
Result Details	Feature			Description Feature Metric			
	source						
		Left End	Right End	Strand	Component Description	Component Metric	Value
İ				F			

		74	133	+			
Hit ID	P-Value	% ident.	Description				Hit Class
EM: CECO9F9:5	7.4e-17		Caenorhabdi	tis elega	ns cosmid C09F9		Trans. Protein
Result Details	Feature		·	Descri	ption	Feature Metric	Value
	source						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		85	152	+			
Hit ID	P-Value	% ident.	Description				Hit Class
☑ EM: CEZK455: 5	8.3e-14	26	26 Caenorhabditis elegans cosmid ZK455				Trans. Protein
Result Details	Feature			Descri	ption	Feature Metric	Value
	source						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		56	134	+			
Hit ID	P-Value	% ident.			Description		Hit Class
M: AC005961:1	1.8e-12	32	32 *** SEQUENCING IN PROGRESS *** Homo sapiens Chromosome 11q12.2 PAC pDJ32L16; HTGS phase 1, 8 unordered pieces.				Trans. Protein
Result Details	Feature		Description Feature Met				Value
1	source						
		Left End	Right End	Strand	Component Description	Component Metric	Value
		112	163	+			
SelectAll 7, 100	eseleci All						

Requested by: Zhelnin, Leonid

Save to Project Submits Lessing





Project Summary

Main Menu	Projects	Query	Seq	Analysis Utiliti	es Logi	n / Logo u	Help	
roject ID: reated: odified: tatus: oday's Date:	<u>orexin</u> OPEN	Created Modifie	d By: Z i ed By: l i User: Z i	rexin receptor hom nelnin, Leonid c_Account, Admin nelnin, Leonid (WR		rch		
Pijoj@citV	Quiens, i	i Fioles	dans.					
Analysis	iools .	Create	Age ni	ry American				
equences (7	of 7 shown	, 7 Active):			<u> </u>			aladan a da
equence ID			Type	Parent	Create	d By	Created	Statu
rexin:g382			DNA	GB:q3823006	Zhelni	•		OPEN
		T Approved Hu =N, genomic :		nomic Sperm Libr sequence.	ary D Ho	mo sapi	ens genor	aic clone
exin:AC00		, , ,	DNA	EM:AC005961	Zhelni	n, L.		OPEN
	ING IN PRO	OGRESS *** H		iens Chromosome		•	32L16; HT	
rexin:fram		61:1	PRT		Zhelni	n, L.		OPEN
exin:newo	rex1.pep:	1	PRT		Zhelni	n, L.	•	OPEN
rexin:G289			PRT	ST:G2897124	Zhelni	n. I		OPEN
REXIN RECE				<u> </u>	J	,		, J. L.
rexin:OX2R			PRT		Zhelni	n. I.		OPEN
		receptor-2		sapiens (human)	01101111	,		
rexin:AF04		receptor z	RNA	EM: AF041243	Zhelni	n. I.		OPEN
		receptor-1 ml			BIICIIII	, 5.		OLDIN
	5 Web 1 1 1 1	MINISTER STREET		Total at at				
in gluatiere	Signification of		Seque.					W
me Agent Re gent Name	sults (10 o	of 166 shown, Started On	166 acti	ve): Description	Status	Method		Job Status
EWOREXIN	+				SAVED		n Agent	Completed
WOREXIN	+		-		SAVED	Protei	n Agent	Completed
WOREXIN	+	:	:		SAVED		n Agent	Completed
WOREXIN	+				SAVED		n Agent	Completed
WOREXIN	+				SAVED		n Agent	Completed
WOREXIN	+				SAVED		n Agent	Completed
WOREXIN	+				SAVED		n Agent	Completed
WOREXIN	+				SAVED		n Agent	Completed
EWOREXIN	+		,		SAVED SAVED		n Agent n Agent	Completed Completed
EWOREXIN	+				× 41 // H' []	シアハアムリリ	. 4/707) 7	LOWDIATA



Query Detail

Seq Analysis Utilities Login / Logout Main Menu Help

Sequence: orexin:AC005961:1

Add Sequence Annotation



Entries w	ith the same Physical Sequence Imagelga:g3859648:1 Newgage:AC005961:1	<u>FASTA</u>			
General Informat	ion About The Entry				
Sequence ID	AC005961				
Version	1				
Sequence Description	*** SEQUENCING IN PROGRESS *** Homo sapiens Chromosome 11q12.2 PAC pDJ32L16; HTGS phase 1, 8 unordered pieces.				
Project	orexin				
Project Name	orexin receptor homology search	orexin receptor homology search			
Molecule	DNA				
Sequence Information					
Length	ength 154647				

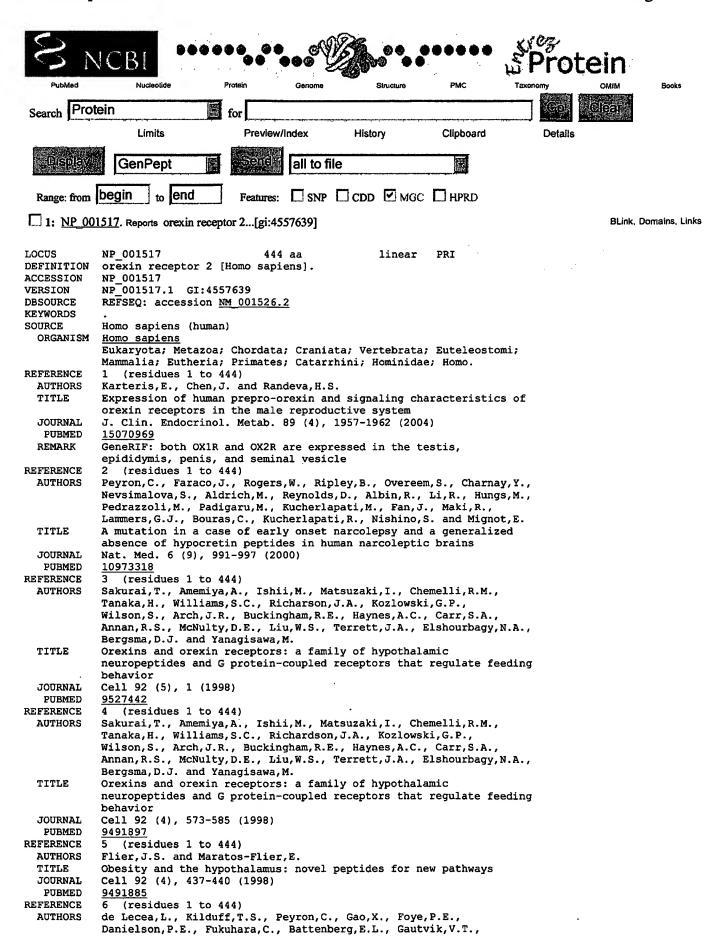


O INCYTE PHARMACEUTICALS, INC.

Exhibit 3

Homology between AC005961 and the orexin-2 receptor

```
Smal
                                                                                 Su
                                                               Reading High
                                                                              Proba
Sequences producing High-scoring Segment Pairs:
                                                                 Frame Score
                                                                              P(N)
AC005961 *** SEQUENCING IN PROGRESS *** Homo sapien..
                                                                   +2
                                                                         103 1.8e-
Score = 103 (48.3 bits), Expect = 1.8e-12, Sum P(5) = 1.8e-12
Identities = 17/52 (32%), Positives = 33/52 (63%), Frame = (+2)
          112 LVVDITETWFFGQSLCKVIPYLQTVSVSVSVLTLSCIALDRWYAICHPLMFK
Query:
                         G + CK + P + + Q + + V
                                           +LT++CIA++R
Sbjct: 48452 LMYDLSYFLTAGAFICKMVPFVQSTAVVTEILTMTCIAVERHQGLVHPFKMK 48607
Score = 88 (41.3 bits), Expect = 1.8e-12, Sum P(5) = 1.8e-12
Identities = 15/54 (27%), Positives = 31/54 (57%), Frame = (+3)
          210 CDERWGGEIYPKMYHICFFLVTYMAPLCLMVLAYLQIFRKLWCRQIPGTSSVVQ 263
Query:
              CEW
                    ++ K+Y ++ ++ PL +M++ Y +I +LW ++ G SV++
Sbjct: 56082 CLEEWTSPVHQKIYTTFILVILFLLPLMVMLILYSKIGYELWIKKRVGDGSVLR 56243
Score = 71 (33.3 bits), Expect = 1.8e-12, Sum P(5) = 1.8e-12
Identities = 15/54 (27%), Positives = 26/54 (48%), Frame = (+1)
          329 VFGMFAHTEDRETVYAWFTFSHWLVYANSAANPIIYNFLSGKFREEFKAAFSCC 382
                      D T+
              + GF
                             F
                                    + ++NS NPI+Y F++ F++
Sbjct: 59380 ISGNFEKEYDDVTIKMIFAIVQIIGFSNSICNPIVYAFMNENFKKNVLSAVCYC 59541
Score = 47 (22.1 bits), Expect = 1.8e-12, Sum P(5) = 1.8e-12
Identities = 9/17 (52%), Positives = 11/17 (64%), Frame = (+2)
Query:
          46 EYLHPKEYEWVLIAGYI 62
              EY H +E W LI+ YI
Sbjct: 29786 EYPHAEE*NWTLISQYI 29836
Score = 46 (21.6 bits), Expect = 1.8e-12, Sum P(5) = 1.8e-12
Identities = 6/8 (75%), Positives = 7/8 (87%), Frame = (+3)
Query:
         381 CCCLGVHH 388
              CCCLG+ H
Sbjct: 148296 CCCLGLEH 148319
```



```
Bartlett, F.S. II, Frankel, W.N., van den Pol, A.N., Bloom, F.E.,
            Gautvik, K.M. and Sutcliffe, J.G.
  TITLE
            The hypocretins: hypothalamus-specific peptides with
            neuroexcitatory activity
  JOURNAL
            Proc. Natl. Acad. Sci. U.S.A. 95 (1), 322-327 (1998)
   PUBMED
            9419374
COMMENT
            REVIEWED REFSEQ: This record has been curated by NCBI staff. The
            reference sequence was derived from AF041245.1.
            Summary: HCRTR2 is a G-protein coupled receptor expressed
            exclusively in the brain. It has 64% identity with HCRTR1. HCRTR2
            binds both orexin A and orexin B neuropeptides. HCRTR2 is involved
            in the central feedback mechanism that regulates feeding behaviour.
FEATURES
                     Location/Qualifiers
     source
                     1..444
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="6p11-q11"
     Protein
                     1..444
                     /product="orexin receptor 2"
                     /note="orexin receptor-2; hypocretin receptor-2"
     CDS
                     1..444
                     /gene="HCRTR2"
                     /coded_by="NM_001526.2:325..1659"
                     /note="go_component: integral to plasma membrane [goid
                     0005887] [evidence TAS] [pmid 9491897];
                     go_function: orexin receptor activity [goid 0016499]
                     [evidence IEA];
                     go_function: rhodopsin-like receptor activity [goid
                     0001584] [evidence IEA];
                     go function: neuropeptide receptor activity [goid 0008188]
                     [evidence TAS] [pmid 9491897];
                     go process: feeding behavior [goid 0007631] [evidence TAS]
                     [pmid 9491897];
                     go_process: synaptic transmission [goid 0007268] [evidence
                     TAS] [pmid 9491897];
                     go process: neuropeptide signaling pathway [goid 0007218]
                     [evidence TAS] [pmid 9491897]"
                     /db xref="GeneID:3062"
                     /db_xref="LocusID:3062"
                     /db_xref="MIM: 602393"
ORIGIN
        1 msgtkledsp pcrnwssase lnetqepfln ptdyddeefl rylwreylhp keyewvliag
       61 yiivfvvali gnvlvcvavw knhhmrtvtn yfivnlslad vlvtitclpa tlvvditetw
      121 ffgqslckvi pylqtvsvsv svltlscial drwyaichpl mfkstakrar nsiviiwivs
      181 ciimipqaiv mecstvfpgl ankttlftvc derwggeiyp kmyhicfflv tymaplclmv
      241 laylqifrkl wcrqipgtss vvqrkwkplq pvsqprgpgq ptksrmsava aeikqirarr
      301 ktarmlmvvl lvfaicylpi silnvlkrvf gmfahtedre tvyawftfsh wlvyansaan
      361 piiynflsgk freefkaafs ccclgvhhrq edrltrgrts tesrkslttq isnfdniskl
      421 seqvvltsis tlpaangagp lqnw
//
```

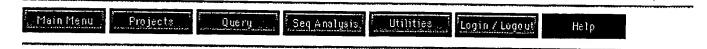
Disclaimer | Write to the Help Desk NCBI | NLM | NIH

Exhibit 4

Description of the 161 bp fragment (NSAN.seq or GA1) and 234 bp fragment (GA2)



Query Detail



Sequence: Newgage:NSAN.seq:1

Add Sequence Annotation



Entries with the same F	hysical Sequence	FASTA					
General Information About The Entry							
Sequence ID	NSAN.seq						
Version	1						
Sequence Description	pcr of g3823006 GA1						
Project	Newgage						
Project Name	Cloning new gene a	gent search sequences					
Molecule	DNA						
Sequence Information							
Length	161						



INCYTE PHARMACEUTICALS, INC.



Retrieve **FASTA Sequences**

Utilities Main Menu Диегц Projects Seq Analysis Login / Logout Help

Project ID: Newgage Project Name: Cloning new gene agent search sequences

Created:

Created By: Zhelnin, Leonid

Modified:

Modified By: Zhelnin, Leonid

Status:

OPEN

Current User: Zhelnin, Leonid (WRITE)

Today's Date:

Num Segs:

>Newgage: NSAN.seq:1 pcr of g3823006 GA1 ${\tt AGTGTACTCATCGCCCTGGCGCTCTTTGGCAATGCTCTGGTGTTCTACGTGGTGACCCGCAGCAAGGCCATGCGCACCGT}$ ${\tt CACCAACATCTTTATCTGCTCCTTGGCGCTCAGTGACCTGCTCATCACCTTCTTCTGCATTCCCGTCACCATGAAGCCGA}$

A

Requested by: Zhelnin, Leonid



O INCYTE PHARMACEUTICALS, INC.



Query Detail

Main Menu Projects Query Seq Analysis Utilities Login / Logout Help

Sequence: Newgage:GA2:1

Add Sequence Annotation



Entries with the	same Physical Sequence <u>FASTA</u>							
General Informati	General Information About The Entry							
Sequence ID	GA2	BA2						
Version	1							
Sequence Description	pcr product of ac ac1 primers of AC00596 genomic chromosome seq in progress							
Project	Newgage							
Project Name	Cloning new gene agent search sequences							
Molecule	DNA							
Sequence Information								
Length	234							



O INCYTE PHARMACEUTICALS, INC.



Retrieve FASTA Sequences

Main Menu Projects Seq Analysis Utilities Query Login / Logout Help

Project ID:

Newgage

Project Name: Cloning new gene agent search sequences

Created:

Created By: Zhelnin, Leonid

Modified:

Modified By: Zhelnin, Leonid

Status:

OPEN

Current User: Zhelnin, Leonid (WRITE)

Today's Date:

Num Seqs:

30

>Newgage:GA2:1 per product of ac ac1 primers of AC00596 genomic chromosome se GTATGATTTATCCTATTTCCTTACTGCAGGTGCTTTCATTTGCAAGATGGTGCCATTTGTCCAGTCTACCGCTGTTGTGA CAGAAATCCTCACTATGACCTGCATTGCTGTGGAAAGGCACCAGGGACTTGTGCATCCTTTTAAAATGAAGTGGCAATAC ACCAACCGAAGGGCTTTCACAATGCTAGGTGAGGCCACTGGATGTGCCAATGGCAGTGTGAATGACATTCTTCA

Requested by: Zhelnin, Leonid



INCYTE PHARMACEUTICALS, INC.

Exhibit 5

Nucleotide sequence comparison of NSAN.seq to SEQ ID NO 1

75 25	150 50	225 75		300		375 125	450 150		525 175		200	675 225		750 250	825 275
TTC	L CH	ပ္ပပ္ ဗ		ပ္ပပ္ပ		CAG	CCT		ATC		TGC	CTG		GGT	ATG
CAG O	GTG V	AAG AAG K		ATT ATT I		GTC v	CAT H		GTC V		ATC	CTC		GAT D	ATG M
GAG E	CTC	AGC AGC S		ည္ ၁၉၈ ၁၉၈		TTT	GTG V		SCA A		CAC H	TTC		999 9	ATT
09 M	ည္တန္	000 000 8		TIC		CCA TIT	CTT GTG		GTG		GAA E	CTC		GTT V	GTC V
ACG T	CTG	ACC ACC		TTC		OTG V	GGA G		TGG CTG W L		AAG K	ATC		AGA R	E SCH
CTG L	AAG K	GTG GTG V		ACC ACC		ATG	CAG Q		TGG W	TM4	GAA E	GTC	TMS	X A	CGA R
AAC	000 4	GTG GTG V		ATC ATC I		AAG ATG K M	CAC H		GTC V		TAT Y	CTT		A.A.G. K	AAA CGA (K R)
CAC H	င်ရှင် န	TAC TAC		CTO CTO D		76C C	AGG R		GTG V		CIA L	ATC		ATA I	AAG K
GAC	66 4 6	TTC		CTG CTG L		ATT TGC I C	GAA AGG E R		GGT G		TTC F	ACC TTC T F		TGG W	ATA GCC AGG AAG AAG I A R K K
000 8	၁၁၁ မ	GTG GTG V		GAC GAC D	TMZ	TTC	GTG V		CTA L		GAC D	ACC		CTT	AGG R
CTG	CTG L	CTG CTG		AGT AGT S		A GCT	GCT		ATG M		TAT Y	ACC		GAA E	000 ₽
CTG L	GAG	GCT GCT		CTC CTC		GGT	ATT I		ACA T		AAA K	TAC		TAT Y	ATA I
ი გ	CC.P	AAT AAT N		908 808 •		999	TGC C		TIC		ATC I	ATC		GGT	TCC AAA S K
TCT	ACC	ည္ဟစ္ ၁ <u>ဗ</u> ္ဗ		TTG TTG L		CTG	ACC T		GCT A		GAG E	AAG K		ATT I	S S
TTC	TAC Y	TTT TTT F		TCC		TGG ¥	ATG M		AGG R		CTT	CAG O		AAA K	ATG M
CAG O	GTC V	CTC CTC L		ပ္အပ္ဆည္သ		AAC N	ACT T	ŀ	CGA R		CAA	CAC H		AGT S	GAA E
GAG E	CTC	900 800		ATC ATC I		GAC	CTC L		AAC N		CAA	GTG V		TAC	AAA K
S S S	000 4	CTG CTG L		rrr rrr F		S H	ATC	TM3	ACC		GTG V	CCT		CTG	GGA G
ACC T	CGA R	ပ္သပ္ ဗင္လ	TMI	ATC ATC I		ATT	GAA	``	TAC Y		E CAC	AGC S		ATT	CAT H
ATT I	CTG	CA TC TTC F		AAC N		AAC	ACA T		CAN O		7 GG	ACC		r rr	ATT
AAC N	CGG R	AGTGTACTCA CTC ATC T' L I F		ACC PACC		GAG CAG	GTG V		TGG W		ATG	TGG W		ATG M	ACT T
CTT L	TAC	GTGT CTC L		GTC GTC V		AAGCCGAA CTC CAG L Q	GTT V		AAG K		ည္မ	GAG		GTG	CGA R
GCG CTT A L	CTG TAC L Y	AGTG' GGC GTG CTC		ACC ACC T		ATG ATG M	GCT GTT		TTT AAA ATG AAG F K M K		TCA CCC S P	TTA GAA GAG L E E		CTT ATG GTG L M V	TCA GIG CIT CGA S V L R
SP. O	GCT A	၁၅၅		ည္ပတ္တ လ		ACC T	ACC		A AA		GGA	TTA L		CIT	GTG V
1 ATG 1 M	76 ATC 26 I	ACC T		ATG CGC ACC GTC ATG CGC ACC GTC		GTC GTC	TCT ACC (S T 1		TTT F		GTA GGA V G	ည်		676 CCT 226 P	TCA
нн	76	151 51		226 76		301	376 126		451 151		526 176	601 201		676 226	751 251

300		975 325		1050 350	1125 375	1200	1275 425	1296 431
AAT N		arcc S		TGC C	A.A.G K	TTG	TCT S	
AGT		AAC		TAT Y	GCA A	AAA K	AAT N	
TAC Y		TCC S		TGT	AAA K	GTC V	GAG E	
GAA E		TTT F		GTT V	AAG K	GAA E	GCT	
ATT		GGA G	2	GCA A	CGG	ATT I	CTG L	
ATG M		ATT	TM7	TCT	ATG M	AAC N	GAA	
ATG M		ATT		TTG L	ATG M	၁ဗ္ဗ	TCT	
GTG ACA GTG GTG CTC TTT GCT GTG TGC TGG GCA CCA TTC CAT GTT GTC ATT GAA TAC AGT AAT V T V V A L F A V C W A P F H V V H M M I E Y S N		TIT GAA AAG GAA TAT GAT GTC ACA ATC AAG ATT TIT GCT ATC GTG CAA ATT ATT GGA TIT TCC AAC TCC F E K E Y D D V T I K M I F A I V Q I I G F S N S		ATC TGT AAT CCC ATT GTC TAT GCA TTT ATG AAT GAA AAC TTC AAA AAA AAA GTT TTG TGT TGT TAT TGC I C N P I S A V C Y C	ADA ACC TTC TCT CCA GCA CAR AGG CAT GGA AAT TCA GGA ATT ACA ATG ATG CGG AAG AAA GCA AAG K T F S P A Q R H G N S G I T M M R K K A K	TCC CTC AGA GAG AAT CCA GTG GAG GAA GGA GAA GCA TTC AGT GAT GGC AAC ATT GAA GTC AAA TTG S L R E N P V E E T K G E A F S D G N I E V K L	AGG R	
GTC		GTG V		AAT N	ATT I	AGT	TTT F	
GTT V		ATC		AAA K	6GA G	TTC	ci L	
CAT H		GCT		AAA K	TCA	GCA A	GCT.	
TTC		TTT F		TTC F	AAT N	GAA E	CIT	
CCA P		ATT		AAC N	GGA G	GGA G	CAT H	
GCA		ATG M		GAA E	CAT H	AAA K	CGA R	
TGG W		AAG K		AAT N	AGG R	ACC	aaa K	
TGC C		ATC		ATG M	CAA	GAA E	CIC	
GTG V		ACA T		TTT F	GCA	GAG E	AAG K	
GCT.	TM6	GTC V		GCA	CCA	GTG V	AAA K	
TTT F	II	GAT D		TAT	TCT	CC.A P	AAG K	TAA *
CTC L		GAT D		STC >	TTC	aat N	GAG E	CAT H
GCT A		TAT Y		ATT	ACC T	GAG E	GAG E	999
GTG V		GAA E		55 4	A A	AGA R	ACA T	AGT
GTG V		AAG K		AAT N	ATA GTA AAT I V N	CIC	CAG Q	GAC D
ACA T		gaa E		TGT C	GTA V	TCC S	GAA E	TTA L
GTG V		TTT E		ATC	ATA	TTT	TGT	CCT
826 276		901 301		976 326	1051 351	1126 376	1201	1276 426
826 GTC 276 V		901 TT 301 F		976 ATC 326 I	1051 AT? 351 I	1126 TT 376 F	1201 TGT GAA CAG ACA GAG AAG AAA AAG CTC AAA CGA CAT CTT GCT CTC TTT AGG TCT GAA CTG GCT GAG AAT TCT 401 C E Q T E E K K K L K R H L A L F R S E L A E N S	1276 CCT TTA GAC AGT GGG CAT TAA 426 P L D S G H *

•

Exhibit 6

Alignment of 234 bp fragment to SEQ ID NO 1

										TCA								
75 25	150 50	225 75		300			375 125	450 150		AGGCCACTGGATGTGCCAATGGCAGTGTGAATGACATTCTTCA	525 175		600 200	675 225		750 250	825 275	
TTC	CTC	9 8 8		SCC 4		CAG	5 0	CCT CCT		TGAA	ATC		7GC C	CTG	1	GGT	ATG M	ĺ
CAG Q	GTG	AAG K		ATT		GTC	Ç <u>1</u>	CAT CAT H		AGTG	GTC V		ATC	CTC I		GAT	ATG M	l
GAG E	CIC	AGC		ညီ ၁			TIL	GTG GTG V		TGGC	GCA A		CAC	TTC		999 9	ATT I	l
095 74	9 8 8	200 M		TTC			e G	CTT		CCAA	GTG V		GAA	CTC		GTT V	GTC	
ACG T	CTG L	ACC		TTC F			org S	668A 668A 6		TGTG	CTG L		AAG K	ATC I		AGA R	GCT.	
CTG	AAG K	GTG V		ACC T		ATG	ATG M	CAG CAG		TGGA	TGG W	TM4	GAA E	GTC V	TMS	AAA K	CGA R	
AAC	GCC A	GTG V		ATC I		AAG	A.A.G	CAC CAC H		CCAC	GTC V		TAT Y	CTT		AAG K	AAA K	
CAC H	၁၉၁	TAC Y		CTC 1		ŢĠĊ	ည် ၁	AGG AGG R		AGG	G GTG V		CTA	ATC I		ATA I	A.A.G.	
GAC	668 G	TTC F		CTG L	TMZ	ATT	ATT I	GAA GAA			661 661		TTC	TTC		TGG W		
C 66	CCG P	GTG V		GAC D	11	TTC	TTC F	GTG GTG V			CTA		GAC D	ACC		CTT	AGG AAG R K	
CTG	CTG	CTG L	1	AGT S			gCI A	GCT GCT			ATG ATG		TAT Y	ACC		GAA E	9 €	
CTG L	GAG E	GCT.		CTC 1		GGT	ត ត្	ATT ATT I			ACA T		AAA K	TAC		TAT Y	ATA I	
090 8	CCA P	AAT N		9 8 8		TGC	ဗ္ဗ ဗ	ညီ ၁ ၁			1117 1117		ATC	ATC		GGT G	AAA K	
TCT	ACC	၁ <u>၅</u> ၅		TTG L		TTAC	ก กับ	ACC ACC	M		GCT A GCT		GAG E	AAG K		ATT I	TCC S	
TTC	TAC	TTT F	. (S S		TTC	TGG ¥	ATG ATG M			AGG AGG R		CTT	CAG		AAA K	ATG M	
CAG O	GTC V	CTC L		၁၅ ၁၅		CTA	N N	ACT ACT T			CGB CGB		CAA	CAC H		AGT S	GAA E	
GAG E	CIC	GCG A		ATC		GTATGATTTATCCTATTTCCTTACTGCA	GAC D	CTC CTC L			AAC N		o Ca	GTG V		TAC Y	AAA K	
CCG P	CCG P	CTG L		TTT		CAT	ည လ	ATC ATC I	TM3		ACC T T		GTG	CCT		CTG L	GGA G	
ACC	CGA R	GCC A	TMI	ATC			ATT	GAA GAA			TAC TAC		CAC H	AGC S		ATT I	CAT H	
ATT I	CTG L	TTC F		AAC N			CAG AAC Q N	GTG ACA GTG ACA V T			CAA		TGG W	ACC T		CTT	ATT I	
CTT AAC L N	000 8	ATC I		ACC T			CAG O	GTG GTG V			7.00 7.00 ¥		ATG M	TGG W		ATG M	ACT	
CIT	TAC	GTG CTC ATC		GTC ACC			ဂ္ဂ ဂ	GTT GTT V			AAG AAG		CCC ATG	GAG E		GTG V	CGA R	
928 4	CTG	GTG V		ACC T			ATG M	GCT GCT			ATG ATG M		ξ S	TTA GAA GAG L E E		ATG M	CTT	
CAG Q	GCT A	၁၅၅	Ŋ.	ATG CGC ACC M R T			GTC ACC ATG	TCT ACC GCT TCT ACC GCT S T A			TTT AAA ATG AAG TGG TTT AAA ATG AAG TGG F K M K W		GTA GGA	TTA L		CCT CTT ATG GTG ATG	TCA GTG CTT CGA ACT ATT S V L R T I	
ATG	ATC	ACC		ATG M				TCT TCT S					GTA	TGC C		CCT	TCA	
нн	76	151 51		226			301	376 126			451 151		526 176	601 201		676	751 251	

300	975 325	1050 350	1125 375	1200	1275 425	1296 431
N N	TCC S	GTC TAT GCA TTT ATG AAT GAA AAC TTC AAA AAA AAT GTT TTG TCT GCA GTT TGT TAT TGC V Y A F M N E N F K K N V L S A V C Y C	AAG K	TTG L	ICT	
S	N N	TAT Y	eca P	AAA K	AAT N	
Y Y	3 TCC	TGT	AAA K	GTC V	GAG E	
¥ 5 🖼	TIT	GTT V	A.A.G	GAA E	GCT A	
i i	TT GGA G TM7	GCA	000 8	ATT I	CIG	
įΣ	I I TA	TCT	ATG M	AAC N	GAA E	
9 ₹ ¥	ATT I	TTG L	ATG M	၁ဗဗ	TCT	
E H	CAA	GTT V	ACA T	GAT D	AGG	
>	org v	AAT N	ATT I	AGT	TTT F	
>	ATC	AAA K	ଜ୍ୟୁ ଜ	TTC	CIC	
H	A A	AAA K	TCA	GCA	GCT A	
Es.	TTT T	TTC	AAT N	GAA	CTT	
а	ATT	AAC N	ଜ୍ୟ ଜ	GGA G	CAT H	
A	GAT GAT GTC ACA ATC AAG ATT TTT GCT ATC GTG CAA ATT ATT GGA TTT TCC AAC TCC D D V T I K M I F A I V Q I I G F S N S TM7	GAA E	CAT H	aat cca gtg gag gaa acc aaa gga gaa gca ttc agt gat ggc aac att gaa gtc aaa ttg n p v e e e t k g e a f s d g n i e v k l	CGA R	
3	AAG K	AAT N	AGG R	ACC	AAA K	
	ATC	ATG	CAA	GAA E	CIC	
} >	ACA T	TTT F	GCA A	GAG E	AAG K	
A TM6	GTC V	GCA A	CCA P	GTG V	AAA K	
F	GAT	Y	TCT	CCA P	AAG	CAT TAA H *
ы	GAT	orc V	TTC	AAT	G.A.G	CAT
8	Y	ATT	ACC T	GAG	GAG	999 9
5 >	E GAA	ပ္သိန္	A.A.	AGA	T.	AGT
5 >	901 TTT GAA AAG GAA TAT 301 F E K E Y	976 ATC TGT AAT CCC ATT 326 I C N P I	1051 ATA GTA AAT AAA ACC TTC TCT CCA GCA CAA AGG CAT GGA AAT TCA GGA ATT ACA ATG ATG CGG AAG AAA GCA AAG 351 I V N K T F S P A Q R H G N S G I T M M R K K A A K	1126 TTT TCC CTC AGA GAG 376 F S L R E	1201 TGT GAA CAG ACA GAG AAG AAA AAG CTC AAA CGA CAT CTT GCT CTC TTT AGG TCT GAA CTG GCT GAG AAT TCT 401 C E Q T E E K K K L K R H L A L F R S E L A E N S	GAC
276 V T V V A L F A V C W A P F H V V H M M I E Y S N TM6	GA.A	TGT C	GTA V	S S	GA.	1276 CCT TTA GAC AGT GGG 426 P L D S G
5 >	TTT	976 ATC 326 I	ATA	TTT	TGT	CCT
276	901 301	976 326	.051 351	.126 376	401	276

Exhibit 7 RACE primers

CROSS REFERENCES:

SU	BJECT _						y by	ر . <u>د</u>	51	RA	CE	السامدي	6	41	133	~-	\$ 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1.	10					: :	• • •		
1 2		1 1	1 CCA	85 TCCTAATA	E 85 ACGACTC	54. ACTATA	-	14.1 27 -4	4.9 3.0 -2	54.1 205.0				48.1 GGGTCA	•	'AGAAC	1 -8.3 ACCAGA	7.9 AGC 25	-41	.9 -1	84.5	-470.9		6AJ	-8	57
3 4		1	CCA	126 ECCTAATA	126 ICGACȚC	54. Actata		17.0 27 - 1	1.7 3.0 -2	56.0 205.0	56.8 -536.4	58	8.6	48.1 <u>GCA</u> GAT	48. Aaagas		-8.3 TGACGO	-7.8 3 <u>16</u> 25		.2 -1	82.9	-471.0		64	-	26
5			1		8	S R	+	+	126	B	-	Al-	IR	(Ne	, j	A	2	-		·		-	·			_
7			1			#	1	4	H	T	-		- 3		_			A		<u> </u>						
8			 		\dashv	\forall	+	+	+	<u> </u>									<u> </u>	-		-				-
9					-	641	- 40	(T	1	1								ó	C	1	+					
10						15	_ [
11				11-		Ţ	1	7		1	1 1		-	11			 	1		-	 	 				
12	_		• •			_		_																		
13	_			xt																						
14				ice F														-0	4	Al	-F	7/				
15	<u> </u>	Se	quen	ice F	ang ما	e:	1 t	0 16	toa	لصي	þr	R	?		6A	1-	85	, D		۱۱ ים ۲	>					
16	- Carte	- d-		Nec 1	1			10		7	20		=	-30		_	4	10		\dashv	50			פש	obe	•
17	_(CAQCT	Cord	70	HEC.	T	GTG CAC	TAC ATG	TCAT AGT	rccc rccc	CCTC	GCQ CGC	CTC GAG	TTT	CCC CCC	AAI TTA	GCI CGA	'CTG	GTG CAC	TTC	TAC	GT CA					•
19			<u> </u>	+		€	₩	L -	A	NSLA	A	L	F	G	N	A	L	V			_V>	·			_	
20	_							60			<		61	41-	12	61	3							\dashv	-	
21					G	GTG.	ACC	CGCZ	GCA	AGGC	CAT	GCG	CAC	CGT	CAC	CAA	CAT	0 CTT	TAT	CTG	.00 CT	-				
22					Ç	$\overline{\mathbf{v}}$	T	g cg₁ R	C 1	Ľλ	M			17	111	× ×		GAA F	ATA	GAC	GA >	+			-	<u> </u>
23					_			6A	TRAI	NSLA	TIO	N O	F G	A1.	TXT	[A]				>	1	_	\dashv		
24					_		_	тμо			.20	•		T30			14	1-1		1	50	7	\dashv	-	_	
25					G	GAA	CCG	GCTC (GAC	TCA	CTGG	ACG	ACT	AGT	GGA	AGA	AGA	ÇGT	AAG	GGC	AGT	CC	•		\dashv	\neg	
26					S	L	A		-5-	NSLA					E	E_	9	7	P	V	T>_			\dashv	-	
27								160									·						+	\neg		
28							AGC	CGAA															+	$\neg \dagger$	\dashv	
29								GCTI P X															7		7	
30									1		İ	-	I],	!		1	-				1	1		1	
31			5	PY	jr	با ا	م			4	0	L		5		R	Ą	- 6	•					一		-
32																										
33		p'	ω	D-S	1																					
34					/																					
ein	NED BY_					1	1		7		***************************************										,				1	<i>i</i>
	NESSED	AND U	NDE	/	OD	RY		<u></u>	<u> </u>	<u>ئ</u>	/			·····			TE . Te .				<u>-</u>		130		-	
											<u> </u>					. ~~	I 154 .									

SUBJECT 2 g3823006. by GA1.txt Aligned Sequence (probe oeigns) Sequence Range: 1 to 481 10 AGCGCGCAAGATGACCTCGTGTCACCGGAGTGTACTCATCGCCCTGGCGC TCGCGCGTTCTACTGGAGCACAGTGGCCTCACATGAGTAGC jmp1 str + AGTGTACTCATCGCCCTGGCG_> 10 GB g382300 AGTGTACTCATCGCCCTGGCGC S'VLI 11 TCTTAGGA GTGTTTACAGTGGTGACCAGCAGCAAGACCATG CACAAATGTCACCACTGGTCGTCGTTCTGGTAC 6AL-Nester 70 GAL-IF ETGACCGCAGCAAGGCCATG> 16 TCTTAGCAAATGCTTTGGTGTTTACAGTGGTGACCAGCAGCAAGACCATG GB q382300 17 18 PATETE TCCTTGGCG TCAGTGACTGCTC 19 GB g382300 CGCACCGTCACCAACATCTZ GCTTGGCAGTGGTTGTAGATAGACC 20 21 jmp1 str + CAGTGACCTGCTC> 22 111111 GB g382300 TCAGTGACTGCTC 23 100 24 CGCACCFTCACCAACATCTTTATCTGCTCCTTGGCGCTCAGTGACCTGCT> 25 CGCACCGTCACCAACATCTATATCTGCTCCTTGGCGCTCAGTGACTGCTC 26 27 180 28 ATCACCTTCTTCTGCATTCCCGTCACCATGCTCCAGAACATTTCCGACAA GB g382300 GAGGEGERAGGAGGTCTTGTAAAGGCTGTT 29 30 ATCACCTTCTTCTGCATTCCCGTCACCATGAAGCCGAA 31 ATCACCTTCTTCTGCATTCCCGTCACCATGCTCCAGAA 32 CATC 33 NOLDS / NILES MI 'SEPT 95 DATE

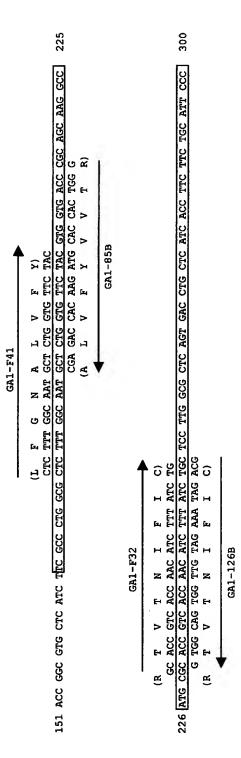
WITNESSED AND UNDERSTOOD BY

Locations of RACE primers on SEQ ID NO 1

1 ATG CAG GCG CTT AAC ATT ACC CCG GAG CAG TTC TCT CGG CTG CTG CGG GAC CAC AAC CTG ACG CGG GAG CAG TTC

75

150 76 ATC GCI CIG TAC CGG CIG CGA CCG CIC GIC TAC ACC CCA GAG CIG CGG GGA CGC GCC AAG CIG GCC CIC GIG CIC



375 525 450 900 750 825 900 975 675 CTC CAG AAC ATT TCC GAC AAC TGG CTG GGG GGT GCT TTC ATT TGC AAG ATG GTG CCA TTT GTC CAG S GGT 451 TTT AAA ATG AAG TGG CAA TAC ACC AAC CGA AGG GCT TTC ACA ATG CTA GGT GTG GTC TGG CTG GTG GCA GTC ATC 526 GTA GGA TCA CCC ATG TGG CAC GTG CAA CTT GAG ATC AAA TAT GAC TTC CTA TAT GAA AAG GAA CAC ATC TGC 601 IGC TTA GAA GAG IGG ACC AGC CCT GIG CAC CAG AAG AIC TAC ACC TCC TTC ATC CTT GIC AIC CTC TIC CTC CTG GTC ATT ATG ATG GCT CTC TTT GCT GTG TGC GCA CCA TTC CAT GTT GTC CAT ATG ATG ATT GAA TAC AGT AAT 901 TTT GAA AAG GAA TAT GAT GAT GTC ACA ATC AAG ATG ATT TTT GCT ATC GTG CAA ATT ATT GGA TTT TCC AAC TCC GTG CAT GGG GAT CTT GTT GTG ACA GAA ATC CTC ACT ATG ACC TGC ATT GCT GTG GAA AGG CAC CAG GGA CTT TGG ATA AAG AAA AGA 751 TCA GIG CIT CGA ACT ATT CAI GGA AAA GAA AIG ICC AAA AIA GCC AGG AAG AAA CGA GCT 676 CCT CTT ATG GTG ATG CTT ATT CTG TAC AGT AAA ATT GGT TAT GAA 376 TCT ACC GCT GTT 826 GTG ACA GTG GTG 301 GTC ACC ATG

0007	1125	1200	1275	1296
ر ا	AAG	TTG	TCT	
IWI	SCA A	AAA	AAT	
15		GTC	GAG	
1	aag aaa	GAA	GCT	
Ş	ອອວ		CTG	
3	ATG	AAC ATT	GAA	
9	ATG	၁၅၅	TCT	
115		GAT	AGG	
AII GIC IAI GCA III AIG AAI GAA AAC IIC AAA AAA AAI GII IIG ICI GCA GII IGI IAI IGC	ATT ACA	AGT GAT	TTT	
Ş	GGA	TTC	CTC	
Ş	TCA	GCA	GCT	
7	AAT.	GAA	CIT	
§	GGA	GGA	CAT	
5	S	AAA	CGA	
Ę	AGG	ACC	AAA	
5	CAA	GAA	CTC	
171	GCA	GAG	AAG	
Ş	SC	GTG	AAA	
141	TCT	CCA	AAG	TAA
ر و	TTC	AAT	GAG	CAT
AII	ACC	GAG	GAG	999
3	AAA	AGA	ACA	AGT
į	AAT	CTC	CAG	GAC
5	GTA	700	GAA	TTA
370 AIC 161 AA1 CCC	ATA GTA AAT AA	TTT	TGT GAA CAG	CCT
9/6	1051	1126 TTT TCC CTC AGA	1201	1276 CCT TTA GAC AGI
	-	•		-

Exhibit 9 5GA1 vs. NPY1 and OX-2

34

BAYER CORP	
SUBJECT GAL - 5 RACE	Product
_	
889 DNA StriderTM 1.2 889	2
Untilled Sequence # 3 -> Genes DMA sequence 347 b.p. astronomonomic chtch/godda linear	3
1 OGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	
82 CMC MAC CITE MOS COM GAS CAS TITC ATTC COTT CITE TAC COM CITE CAA COM CITE ATC PC TO COM CAS CITÉ COM C	
157 COC COC AND CITG GOC CITC GITC CITC ACC COC CITC ACC CITC GOC CITC GOC CITC GOC ANT GOT CITC GOC GOC ANT GOT CITC GOC GOC ACC CITC GOC ANT GOT CITC GOC GOC ACC CITC GOC ANT GOT ACC GOC GOC GOC GOC GOC GOC GOC GOC GOC G	
232 THC GTG GTG ACC GCC ACC ANG GCC ATG GCC ACC GTC ACC ANG ATC TIT ATC TCC TCC TTG ACC CTC ACT GAC CC 307 CTC ATC ACC TTC TTC TCC ATT CCC GTC ACC ATC AAA GCC AA 301 CTC ATC ACC TTC TTC TCC ATT CCC GTC ACC ATC AAA GCC AA 301 CTC ATC ACC TCC ATC ACC TCC ATC ACC ATC AAA GCC AA	347
- 93 L I T F F C I P V T N K P	105
	9
5GA1 Aligned Sequences Formatted Alig	
	<u> 11 </u>
	1 12
	13
ClustalW Formatte	d Alignments 14
;	15
	16 ,
5GA1 MOMENTAL NITTOP FOR SECTION	20 30
5GA1 O A LNITPE OFSRIN NPY-1R.pep NSSTLFSQVENHS VI	LRDHNLTRE F LYR 17 HSNFSEKNAS L FEN 18
	,
	19
40	50 60 20
5GA1 LRPLVYTPE GRAI NPY-1R.pep D DCH DLAM	
NPY-1R.pep D D C H	I F A Y G A Y I I G Z 22
1.	
5GA1 NPY-1R.pep S L A L V F Y V V T R S L A L I I L L R Q I I	80 90 24
NPV-1P con S S I A T J J J J J J J J J J J J J J J J J J	A CRITICAL CS BAL 25
W 1-111.pep 3 No. 1 A LEGIT I LEGIC LEGI	L SV N S F S C C C C C C C C C C C C C C C C C
'	
100	110 120 27 28 Y T L M D H W V F G E A M C K L 29
5GA1 DEFFFORVVMKI	P ^F 28
NPY-1R.pep D & VA I M C D F F F V Y	YTLMDHWVFGEAMCKL 29
	30
	31
	32
	33
	

REVIOU DE A REVIOU DE / NILES, MI ' SEPT. 95 _____ DATE WITNESSED AND UNDERSTOOD BY Jun Withou DATE

SUBJECT

SAL PACE

													T
	Orexin	2.pep	Aligned	l Sequence	es Forma	tteä	314 mg		_				
					JO I OIMO	rcea	WIIGI	ment	8				
<u> </u>													
												i	
			Clusta	lW Fo	rmatte	ed A	A lio	nma	nte	!			
						•	5	******	- II ()	,			
_				40									
	Orexin 2	.pep	SGTK	10 E D S P	CRN	wss	ا ی	20 100 r 15	Mc I		In Ma	30	
_	5GA1	W	QALN	E D S P T	PEQF	SRL	LR	DН	L	R	0	I N	\prod
									2411	· · · · · · · · · · · · · · · · · · ·		-	П
_				. 40			£	50				60	
_	Orexin 2	.pep P	TDYD	DEEF	YW	REY	LH	K	EYE	wv	T.	A G	- 1
	·JUAI	L	I K		⊠ P V	YTP	EL	G	RAK	LA			
, s		•										210 by	67
	Orexin 2.	pep y	ii ii v	70 V V V T	V2 558 v 657	- ·	_	0 	_		Sev San	90	2 4
<u> </u>	5GA1	G	Varius II	ALAIF	V I	FY	A V	W K N	H	MATERIAL SERVICES	44 - 44 . 44 - 44 .		
					5 A 7	. 403	The self-or		. F.W.	EXTER	a diamental		
				100			11	n				400	
	Orexin 2.	pep Y	NV	I SI I WIN	V i V i	IT		A	lv v	DΙ	ТЕ	120 T W	
	5GA1	I	C S			F F	DE IN	V N	K P			4	
23	+++-					1 1	4-4-			$\perp \perp$		Γ	
24		 					1-1-					1	
25	- - - -	g1 14758820	Pauropantida (<u></u>]	1		\perp		1	
26	++		(AF119815) G-1 Length = 522	protein-coupled protein-coupled re	receptor gb AA ceptor (Homo sa	w22047.1 apiens]		44		11		\sqcup	\bot
27	++	Score = 15; Identities	2 (73.0 bits), = 27/55 (49%),	Expect = 1.8e-13, Positives = 42/5	P = 1.8e-13 5 (76%)			1				 	
28	- - 	Query: 49	VLTGVLIFALALE	GNALVFYVVTRSKAMRT	VTNIPICSLALSDLI	7. EC.D.M.				-		$\downarrow \downarrow \downarrow$	
29	- - 		LIISYFLIFFLCM	GNTVVCFIVMRNKHMHT	VTNLFILNLAISDLI	LVGIFCMPITI	205	+		-		1	
30		>SP:NY2R_PIG	NEUROPEPTIDE Y	RECEPTOR TYPE 2			-06			11		11	
31	s	mith-Waterman	n score: 181;	45.1006 146.6161	40	50	60 _			1 1		$\downarrow \downarrow$	
32	i	gal		SGPGHTTPRGELAPDSEP	ELKDSTKLIEVOII	LILAYCSIIL	LGVVG	+		- -		+	
33		SP:NY2R_PIG	10 20	30 4 on 90	100	60		+-		+	_		
34	!	gal	NALVFYVVTRSKA	MRTVTNIFICSLALSDLL	.ITFFCIPVTMKP	MGEWKMGPVL	CHLVP	+		 		 	
		SP:NY2R_PIG	70 80	90 10	,,,								
SIGNE) BY	SP:NY2R_PIG	YAQGLAVQVSTIT	LTVIALDRHRCIVYHLES	SKISKRISFLIIGLE 60 170	180			*	<u>, </u>	-·.· -	"LEE	th .∓m:
	SSED AND UN	DERSTO	OD BY	Juni 1	7/t		_ DAT	도 =				<u></u>	-
	REFERENCE			S. DATILI	1/////		<i>U</i> AT	5 ,		- 			-

Alignment of 5GA1 with SEQ ID NO 2

25	20	75	100	125	150	175	200	225	250	275	300	325	350	375	400	425	431
[24 [24	HH	4 4	<u>a</u> a	ø	Q,	н	ပ	ы	ტ	×	z	တ	ပ	×	ы	w	
aa	>>	* *	нн	>	Ħ	>	н	ы	Q	×	တ	z	*	æ	×	z	
10 to	ս리	တတ	ပြပ	[24	>	4	×	Ēų	ပ	н	>	တ	υ	×	>	闰	
44	44	~ ~	[E4 [E4	Д	ы	>	M	ы	>	>	阳	[24	>	×	团	æ	
H H	a a	타타	[21 Ex	Δ	_o	ы	M	н	æ	æ	н	ტ	Æ	α	н	H	
111	× ×	> >	E E	Σ	α	Z	E	>	M.	ρ¢	Σ	н	Ø	×	z	回	
22	AA	 	нн	×	×	>	> +	ы	×	×	Σ	н	д	×	ဖ	တ	
# #	82 82	N N	14 H	ပ	œ	>	н	н	н	×	×	œ	>	E+	Ω	~	
۵۵	ن ق	Eu Eu	H H	н	闰	_O	ĹΉ	(Eu	3	×	>	>	z	н	Ø	Ēŧ	
04 04	дη	>	ДΩ	Ē	>	H	Д	E4	ы	ρζ	>	н	×	Ŋ	Ēu	ы	
44	44	111	s s	4	4	Ξ	-	•· E+	ы	- α	H	«	M	ω ω	4	~	
1 1		1 1										Ei -	E4				
그리	ខេធ	AA	17 17	. U	н	H	×	X	×	н	Įъ			Z	(2)	П	
~ ~	<u>а</u> , а,	ZZ	AA	Ŋ	υ	Œ	H	н	O	×	Д	H	Z	Ø	o	×	
တ တ	H H	ပ ပ	H H	ы	H	æ	M	×	н	Ø	æ	Σ	ഥ	×	×	æ	
Eu Eu	× ×	նանա	တ တ	3	Σ	æ	,1	a	×	Σ	3	×	z	œ	H	×	
αa	> >	ㅋㅋ	ပပ	z	H	æ	œ	#	တ	ធ	ပ	н	×	œ	回	н	
ωы	H H	44	нн	Ω	н	z	œ	>	*	×	>	E	드	Ø	딥	×	
0.0	а а	ㅂㅂ	Eu Eu	တ	н	H	>	ф	н	ဖ	Ø	>	æ	Д	>	M	
H H	~ ~	44	нн	н	回	>	Ħ	w	н	×	Ē	Ω	> 1	w	щ	×	*
нн	교교	E4 E4	22	Z	H	α	:	H	H	н	ы	Ω	>	Œ	z	덟	Ħ
zz	84 84	нн	E1 E1	α	>	3	Σ	3	E	H	Ø	>	н	H	函	闰	დ
ㅁㅁ	× ×	нн	> >	ᆸ	>	×	ρι	ы	>	œ	>	闰	ω	×	c c	E	တ
44	4 4	>>	E+ E+	[<u>z z</u>]	æ	Σ	တ	ы	×	н	>	×	z	z	ы	a	Ω
aa	e e	ს ს	K K	E+ E+	EH	×	U	н	13	>	H	⊠ .	ပ	>	w	ഥ	H
ΣΣ	нн	E E	ΣΣ	>>	S	Ēυ	>	υ	ρı	w	>	[Es	н	н	[±4	U	Дı
	26	51	94	101	126	151	176	201	226	251	276	301	326	351	376	401	426 P

-

Alignment of 5GA1 nucleotide sequence with SEQ ID NO 1

30	105	75
CA	TC	밁
ACAG	T SK:	AG T
GCGC	AG C	AG C
GGGA	ე ეე	9
rccc	၁	9
	IG A	IG A
ညည	S S	S
SGTG	CA	N S
•	ပ္	S
	9	<i>5</i> 5
	ပ္ပ	5
	ភ្ជ	10 E
	ည် ပ	T C
	5 1	5
	J.	J.
	E	TI
	2	CAC
	8	GAG
	ဗ္ဗ	
	P S S	S
	ATT	ATT
	AAC	AAC
	IG CAG GCG CIT AAC ATT ACC CCG GAG CAG TTC TCT CGG CTG CTG CGG GAC CAC AAC CTG ACG CGG GAG CAG TTC	IG CAG GCG CIT AAC AIT ACC CCG GAG CAG IIC ICI CGG CIG CIG CGG GAC CAC AAC CIG ACG CGG GAG CAG IIC
	ပ္ပြ	939
	CAG	CAG
	ပ္ည	9

105 75	25	180	150	20
TT AAC ATT ACC CCG GAG CAG TTC TCT CGG CTG CTG CGG GAC CAC AAC CTG ACG CGG GAG CAG TTC TT AAC ATT ACC CCG GAG CAG TTC TCT CGG CTG CTG CGG GAC CAC AAC CTG ACG CGG GAG CAG TTC	Ĺų	TAC CGG CTG CGA CCG CTC GTC TAC ACC CCA GAG CTG CCG GGA CGC GCC AAG CTG GCC CTC GTG CTC	CTC	u
CAG	a	GTG	GIG	>
GAG	ы	CIC	5	ч
ည် ည	æ	ည္ဟ	ပ္ပ	Æ
90g 90g	E	c_{TG}	SE	ı
CTG	ы	AAG	AAG	×
AAC AAC	z	ည္ဟ	ပ္ပ	Æ
CAC	H	ည္သ	ည	œ
SAC SAC	Ω	GGA	GGA	G
999	æ	၁၁	ဥ္ပဌ	Д
CTG	ı	CTG	CTG	μŦ
5 TO 5	н Н	GAG	GAG	回
ဗ္ဗ ဗ္ဗ	~	SGA	CCA	д
រុះ រុះ	တ	ACC	ACC	Ŀ
DE E	[b ₄	TAC	TAC	LVYT
CAG CAG	æ	STC	GIC	>
GAG	回	STO	CIC	ч
ဗ္ဗ ဗ္ဗ	Д,	ည	ည	Ā
ACC	£	SGA	CGA	æ
ATT	н	CTG	CTG	1
AAC AAC	Z	ဗ္ဗ	999	24
CTT	7	TAC	TAC	×
CAG GCG	Æ	CTG	CTG	 -1
	a	P.	GCT	 4
ATG	Σ	ATC	ATC	I
31	H	106	9/	26

²⁵⁵ 225 75 GGC AAT GGC AAT GG N N 181 ACC GGC GTG CTC ATC TTC GCC CTG GCG CTC TTT
151 ACC GGC GTG CTC ATC TTC GCC CTG GCG CTC TTT
51 T G V L I F A L A L F

330	300	100
သည	SIC ACC AAC AIC ITI AIC IGC ICC IIG GCG CIC AGI GAC CIG CIC AIC ACC IIC IIC IGC AII CCC	Δ,
ATT	ATT	H
TGC	TGC	ບ
TTC	TTC	ĺΨ
TTC	TTC	Ĺ
ACC	S	H
ATC	ATC	н
CTC	CTC	П
CTG	CTG	'n
GAC	GAC	Ω
AGT	AGT	SLALSD
CIC	CTC	ы
929	ည	A
TTG	TTG	ı
100	TCC	S
TGC	TGC	NIFICS
ATIC	ATC	I
TTT	TTT	Ŀ
ATC	ATC	H
AAC	AAC	z
ACC	ACC	F
GTC	0	>
S ACC GI	ACC.	H
ညည	26 ATG CGC	m m
ATG	ATG	Σ
256	226	16

³⁴⁷ 375 125 IGTC ACC ATGRAGCCGAA IGTC ACC ATG CTC CAG AAC ATT TCC GAC AAC TGG CTG GGG GGT GCT TTC ATT TGC AAG ATG GTG CCA TTT GTC CAG V T M L Q N I S D N W L G G A F I C K M V P F V Q

450	150
CCT	ы
CAT	I
GTG	>
\ CTT	႕
GGA	ტ
CAG	a
CAC	Ħ
AGG	œ
GAA	臼
GTG	>
GCT GTG	Ø
ATT	н
TGC ATT	ပ
ACC	E
	Σ
ACT	E
ST	ы
ATC	н
GAA	
, ACA	E
GTG	>
GTT	>
ACC GCT	A
ACC	E
TCT	Ø
376 TCT	126 S

525	175
ATC	H
GTC	>
GCA	æ
GTG	>
CTG GTG	'n
\mathbf{TGG}	3
GTC	>
GTG	>
GGT	G
CTA	ᆸ
ATG	Σ
ACA	H
TTC	ഥ
GCT	Æ
AGG	×
4 50	~
AAC	z
ACC	H
TAC	>+
TGG CAA	a
ŢĞĞ	3
AAG	×
ATG	Σ
TTT AAA ATG	×
ŢŢŢ	[z.
51 1	151
4	H

009	200
TGC	ບ
CCC AIG TGG CAC GTG CAA CAA CTT GAG ATC AAA TAT GAC TTC CTA TAT GAA AAG GAA CAC ATC TGC	н
CAC	H
GAA	闰
AAG	×
GAA	E)
TAT	~
CIA	7
TIC	[z.
GAC	٥
TAT	>
AAA	×
ATC	-
GAG	Œ
CIT	_
CAA	c
CAA	c
GTG	٥
CAC	=
TG	3
ATG	>
ပ္ပ	ρ
Ę	U
GG.	į.
526 GTA GGA	•
526	176 17

675	225
CTG	ᆟ
5 E	н
TTC CTC	(Fr)
SE	ы
ATC	н
ATC CIT GIC	>
CTT	ы
ATC	н
ACC TTC	Ē
ACC	YTTF
ACC	H
S AAG ATC TAC ACC	×
ATC	н
AAG	ĸ
CAG	a
GTG CAC	H
GTG	>
CCI	Д
AGC	ß
GAG TGG ACC	H
TGG	×
GAG	回
GAA	
TTA	н
TGC	U
601	

⁸²⁵ 275 ATG M ATG M GTC ATT GCT A AAA CGA C AAA GAA ATG TCC AAA ATA GCC AGG AAG AAG K E M S K I A R K K **4**55 5 CAT H ATT I CGA ACT R CTT 751 TCA (251 S

³⁰⁰ AAT N GTG TGC TGG GCA CCA TTC CAT GTT GTC CAT ATG ATT GAA TAC AGT V C W A R P F H V V H M M I E Y S CIC III GCI GTG GTG GCT V V A

975	1050	1125	1200	1275	1296
325	350	375	400	425	432
GAA TAT GAT GAT GTC ACA ATC AAG ATT TTT GCT ATC GTG CAA ATT AIT GGA TTT TCC AAC TCC E Y D D V T I K M I F A I V Q I I G F S N S	CCC ATT GTC TAT GCA TTT ATG AAT GAA AAA AAA GATT TTG TAT TGC P I V Y A F M N E N F K K N V L S A V C Y C	AAA ACC TTC TCT CCA GCA CAA GGA AAT TCA GGA ATT ACA ATG ATG CGG AAG AAA GCA AAG K T F S P A Q R H G N S G I T M M R K K A K	aga gag aat cca gtg gag gaa acc aaa gga gaa gca ttc agt gat ggc aac att gaa gtc aaa ttg r e n p v e e t t k g e. a f s d g n i e v k l	TCT	
AAC	TAT	GCA	AAA	AAT	
N	Y	A	K	N	
acc	TGT	AAA	GTC	GAG	
S	C	K	V	E	
TTT	GTT	AAG	GA.A	GCT GAG AAT TCT	
F	V	K	E	A E N S	
6GA 6	A GCP	CGG 78	ATT	CTG L	
ATT I	S	ATG M	N N	GAA E	
ATT I	TTG L	ATG M	၁ ၀ ၅	TCT	
CAA	GTT V	ACA T	GAT D	AGG R	
GTG V	AAT N	ATT	AGT	TTT F	
ATC	A A A A	GGA G	TTC	r L	
GCT A	A A A A	S	GCA A	GCT A	
TTT F	TTC	AAT N	GAA E	CTT	
ATT I	AAC N	ogA G	66 8	CAT H	
ATG	GAA	CAT	AAA	ACA GAG GAG AAG AAA AAG CTC AAA CGA CAT CTT GCT CTC TTT AGG TCT GAA CTG	
M	E	H	K	T E E K K K L K R H L A L F R S E L	
A.A.G	AAT	AGG	ACC	AAA	
K	N	R	T	K	
ATC	ATG M	CAA	GAA E	CTC 1	
ACA	TTT	GCA	GAG	AAG	
T	F	A	E	K	
GTC	GCA	CCA	GTG	AAA	
V	A	P	V	K	
GAT	TAT	TCT	CCA	AAG	TAA
D	Y	S	P	K	*
GAT	GTC	TTC	AAT	GAG	CAT
D	V	F	N	E	H
TAT	ATT	ACC	GAG	GAG	ე
Y	I		E	E	ე
GAA	CCC	AAA	AGA	ACA	AGT GGG CAT
E	P	K	R	T	S G H
AAG	AAT	AAT	CIC	CAG	GAC
K	N	N		O	D
GAA	TGT	GTA	သ	GAA	TTA
E		V	သ	E	L
TTT F	ATC I	ATA I	TTT F	TGT	CCT
901 301	976 326	1051 351	1126 376	1201	1276

Exhibit 12 3' RACE primers

31 RACE

1	
2	Translated Sequence New Page 1
3	Translated Sequence 6A1 F 41
5	
6	Range: 1 to 161
7	10 20 30 40 50 Mention 6A 1 AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCATCGCCCTGGCGCATGCTCTGGTGTTCTAGGT AGTGTACTCTGGCGATGCTCTGGTGTTCTAGGT AGTGTACTCTGGTGTTCTAGGTATGCTCTGGTGTTCTAGGT AGTGTACTCTGGTGTTCTAGGTATGCTCTGGTGTTCTAGGT AGTGTACTCTGGTGTTCTAGGTATGCTCTGGTGTTCTAGGT AGTGTACTCTGGTGTTCTAGGTATGCTCTGGTGTTCTAGGT AGTGTACTGCTCTGGTGTTCTAGGTATGCTCTGGTGTTCTAGGT AGTGTACTGTAGGTATGCTCTGGTGTTCTAGGTTTAGGTTTAGGTTTAGGTTTAGGTTTAGGTTTAGGTTTAGGTTTAGGTTTAGGTTAGGTTTAGGTT
8	TCACATGAGTAGCGGGACCGCGAGAAACCGTTACGAGACCACAAGATGCA
9	S V L I A L A L F G N A L V F Y V>TRANSLATION OF GA1.TXT [A]>
10	Nexted 6A1
11	GETGACCCGCAGCAAGGCCATGGCACCGTCACCAACATCTTTATCTGCT (AF 3)
12	CCACTGGGCGTCGTTCCGGTACGCGTGGCAGTGGTTGTAGAAATAGACGA V T R S K A M R T V T N I F I C>
13	
14	110 120 130 140 150
15	CCTTGGCGCTCAGTGACCTGCTCATCACCTTCTTCTGCATTCCCGTCACC GGAACCGCGAGTCACTGGACGAGTAGTGGAAGAAGACGTAAGGGCAGTGG To order
16	TRANSLATION OF GA1.TXT [A]
17	
18	160 3 Race —
19	TACTTCGGCTT M K P X>
20	
21	
22	41 22 188 167 53.9 78.0 3.5 56.7 60.4 56.8 48.1 48.1 -3.6 -8.3
23	CTCTTTGCCAATGCTCTGGTGTTCTAC 27 -43.6 -199.5 -515.1 * CCATCCTAATACGACTCACTATAGGGC 27 -43.0 -205.0 -536.4
24	281-741 (Blan Browning 1/19105)
25	
26	32 74 188 115 52.2 75.4 1.7 54.9 58.6 56.8 48.0 48.1 -4.4 -8.3
27	GCACCGTCACCAACATCTTTATCTG 25 -40.2 -182.9 -471.0 *
28	GAS-F32 (Brian Farmerical / /q(05)
29	
30	
31	
32	
33	
34	
	NED BYDATE
WITI	NESSED AND UNDERSTOOD BY

SUBJECT	Touch	Dow	n ·	PCI	٤	for	68	41	R	AE				
		-			_\				<u></u>					
(15) 5W CDNA	templa	10	16	leans		701	<u> </u>	1.7	7=	Tat				
2/ 1	Tempen		-(-			(5)	91	~	<i>/</i>	11-				
	06				_/_	4		1		2				
5 LO LOX CON	A Buffe	72					'			3				
(3) 1/Le 2/N/1/	lixlion	H)								4				
(3) 140 Ads		en m	1x/	SOX					-	5				
		11-	77 (/ 			-						
U3 1 0	100		1 2			<u> </u>	_	∤ · ↓		6				
rogram 1 (prefe	nac	77000	+ 3	12 P	AP	1		1	•	7				
PF.DNA.Thom	T'T T	7 I _m > /0°C):					i	í		8				
			– (or	GeneAm hot-lid the	p Syste	ms 2400	/9600			9				
1,00,000	2	1-1-1	_ `	• 94°C fo	r 30 se	C				10				
Klard 1 B12	6		_	 5 cycles 	s:					111				
9 2 F 3	2 /			72°C	5 sec 4 min	٥				12				
1/3/14				 5 cycles 	: :				l	13				
0 9 8126				70°C	5 sec 4 min	>			ľ	14				
I DOG ALL I I I I			_	• 20–25 c	ycles: 5 sec				F	15				
1 1 1 F32				68°C	4 min	•	•		<i> </i>	16				
6 F41				г.т. Т.	[".]	1 1	1		. ج <u>ئ</u> ــ	17				
				 		-+-+				18				
RODUCT: Human	RODUCT: Human Heart													
Marathor	n-Ready™ d	DNA						;		19				
\TALOG #: 7404-1										20				
17404-1		DESCRIPT		DMA io	biek e.	ا المائم،				21				
 -		Marathon-Restranded cD	DNA w	hich has	nign-qu haan	lality, di	ouble-			22				
Τ#: <u>9010719</u>		Marathon™	Adapto	or and is	ready	for use	as a			23				
		template in 5	5' and 3	3' Maratho	n RAC	E reaction	ns. In			24				
		many cases	the 1	full-length	cDNA	can th	en be							
STORAGE BUFFER:		obtained by cloning. En	onap ena-to	-end amp material	is pro	n or sta	ndard	Ü		25				
10 mM Tricine-KOH (pH 9.2)		50-μl Marath	on RAC	CE reactio	ns.	vided i	01 30			26				
·										27				
STORAGE CONDITIONS:		CONCENTR	ATION	i: ≈ 0.1	ng/μl					28				
Avoid multiple freeze/thaw o	ycles.	POLY A' R	NA S	OURCE:						29				
		Normal, wh	nole h	earts po	oled f	rom 3	male			30				
SHELF LIFE: year from date of recei		Caucasians,	ages 28	3-47; caus	e of de	ath: traur	na			11				
proper storage conditions		No further RN	lΔ eour	o info	tion in n	wailahla		÷		2				
			in auult ! !	Simonn e I I	uon is a	valiadie.		i, i		33				
			$\neg \vdash \vdash$		++	++	+	-	+	34				
							#			58, MI / SEPT. 95				
SIGNED BY			<u>へ</u>	- 1	DATE _		•			.o, mi ' SEP1. 95				
WITNESSED AND UNDERSTOOL	BY Ju	in With	ź		DATE _	14-1		. :	- , :					
	•													

3'-RACE #2 is assembled from LZLZ11 and LZLZ21

SUE	BJECT huma GAL project. Full length 5't3' ends	-
	heis sequence was deposited in the Life Ison of unknown protein and called GA1, MO.	1 62 3 4
/ 6A=	J-corward 3'Race#2clone(8.27) Sequencherm "gal 3'65"	5
5 GA1 consensus	atg cag gcg gtt aac att acc ccg gag cag ttc tct cgg ctg ctg cgg gac cac aac ctg acg cgg gag cag ttc atc gct ctg tac cgg ctg cga ccg ctc gtc.	8 9
5'GA1 consensus. #106	tac acc cca gag ctg ccg gga cgc gcc aag ctg gcc ctc gtg ctc acc ggc gtg ctc atc ttc gcc ctg gcg ctc ttt ggc aat gct ctg gtg ttc tac gtg gtg tac acc cca gag ctg ccg gga cgc gcc aag ctg gcc ctc gtg ctc acc ggc gtg ctc atc ttc gcc ctg gcg ctc ttt ggc aat gct ctg gtg ttc tac gtg gtg	10 11
5'GA1 consensus 45_1-77-711 4211	ace cgc agc aag gcc atg cgc acc gtc acc aac atc ttt atc tgc tcc ttg gcg ctc agt gac ctg ctc atc acc ttc ttc tgc att ccc gtc acc atg gc acc gtc acc aac atc ttt atc tgc tcc ttg gcg ctc agt gac ctg ctc atc acc ttc ttc tgc att ccc gtc acc atg acc cgc agc aag gcc atg cgc acc gtc acc atc atc ttt ttc tgc att ccc gtc acc atg ctc cag acc cgc agc aag gcc atg cgc acc gtc acc atc atc ttt ttc tgc att ccc gtc acc atg ctc cag	12 13 14
45_121211 46_121221 #316	aac att too gad aac tgg etg ggg ggt get tie att tge aag atg gtg coa tit gte cag tot ace get git gtg aca gaa atc cie act atg ace tgc att gct.	15 16 17
45 J.21.201 46 J.21.221 #421	S D N W L G G A F I C K M V P F V Q S T A V V T E I L T M T C I A gtg gaa agg cac cag gga ctt gtg cat cct ttt aaa atg aag tgg caa tac acc aac cga agg gct ttc aca atg cta ggt gtg gtc tgg ctg gtg gca gtc atc gtg gaa agg cAc cag gga ctt gtg cat cct ttt aaa atg aag tgg caa tac Acc aac cga agg gct ttc Aca atg cta ggt gtg gtc tgg ctg gtg gca gtc atc	18 19
45_] <i>2</i> 7.27.1 46_].27.22.1 #526	V E R H Q G L V H P F K M K W Q Y T N R R A F T M L G V V W T V 2 T N R R A F T M L G V V W T N R R A F T M R R A F T M L G V V W T N R R A F T M L G V V W T N R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R A F T M R R R A F T M R R R A F T M R R R R R R R R R R R R R R R R R R	20 21 22
45_ <i>L2</i> 7.211 46_J.27.221	V G S P M W H. V Q Q L E I K Y D P L Y E K E H I C C L E E; W T S P V H	23
2221	cag aag atc tac acc acc ttc atc ctt gtc atc ctc ttc ctc ctg cct ctt atg gtg atg ctt att ctg tac agt aaa att ggt tat gaa ctt tgg ata aag aaa Q K I Y T T F I L V I L F L L P L M V M L I L Y S K I G Y R L W I K K aga gtt ggg gat ggt tca gtg gt ggt ggt tca gtg gt ggt ggt tca gtg ctc ctc atc ctc ctc ctc ctc ctc ctc ctc	25 26 27
, k11	aga gtt ggg gat ggt tca gtg ctt cga act att cat gga aaa gaa atg tcc aaa ata gcc agg aag aag aag aag ag gct gtc att atg atg gtg aca gtg gtg gct R V G D G S V L R T I H G K E M S K I A R K K R A V I M M V T V V A Ctc ttt gct gra tgg tgg tgg tgg tgg tgg tgg tgg tgg tg	28 29
841	ctc tit gct gtg tgc tgg gca cca ttc cat git gtc cat atg atg att gas tac agt aat tit gas aag gas tat gat gat gtc aca atc aag atg att tit gct ctc tit gct gtg tgc tgg gca cca ttc cat git gtc cat atg atg att gas tac agt aat tit gas aag gas tat gat gat gtc aca atc aag atg att tit gct L F A V C M A P F H V V H A M L E Y S N F E K B Y D D V T I K M L R A	30 31 32
	atc gtg caa att att gga ttt tcc aac tcc atc tgt aat ccc att gtc tac gca ttt atg aat gaa aac ttc gaa aaa aaa aaa aaa aaa aaa aaa aaa aa	33
	cta gac ctg ccc ggg cgg ccc gct cga gcc cta tag tga gta tcg gat ccc cgg gta ccg agc tcg aat tca ctg gcc gtc gtt tta caa ag CL D L P G R P A R A L V S D P R V P S S N S L	MI / SEPT.

Alignment of 3'-RACE #2 with SEQ ID NO 1 and SEQ ID NO 2

75	225	300	146 375	221 450	296 525	371	446	521 750	596 825	671 900	746 975	807 1050	1125
1 ATG CAG GCG CIT AAC AIT ACC CCG GAG CAG TIC ICI CGG CIG CIG GGG GAC CAC AAC CIG ACG CGG GAG CAG TIC 76 AIC GCI CIG IAC CGG CIG CGA CCG CIC GIC IAC ACC CCA GAG CIG CCG GGA CGC GCC AAG CIG GCC CIC GIG CIC	151 ACC GGC GTG CTC ATC TTC GCC CTG GCG CTC TTT GGC AAT GCT CTG GTG TTC TAC GTG GTG ACC CGC AGG AAG GCC 2	GC ACC GTC ACC AAC ATC TIT ATC TGC TCC TTG GCG CTC AGT GAC CTG CTC ATC ACC TTC TTC TGC ATT CCC CGC ACC GCC ACC ACC ATC TTC TGC ATT CCC	12 GTC ACC ATG CTC CAG AAC ATT TCC GAC AAC TGG CTG GGG GGT GCT TTC ATT TGC AAG ATG GTG CCA TTT GTC CAG 301 GTC ACC ATG CTC CAG AAC ATT TCC GAC AAC TGG CTG CTG GGG GGT GCT TTC ATT TGC AAG ATG GTG CCA TTT GTC CAG 3	147 TTCT ACC GCT GTT GTG ACA GAA ATC CTC ACT ATG ACC TGC ATT GCT GTG GAA AGG CAC CAG GGA CTT GTG CAT CCT 376 TCT ACC GCT GTT GTG ACA ATC CTC ACT ATG ACC TGC ATT GCT GTG GAA AGG CAC CAG GGA CTT GTG CAT CCT 4	222 TIT AAA ATG AAG TGG CAA TAC ACC AAC CGA AGG GCT TTC ACA ATG CTA GGT GTG GTC TGG CTG GTG GCA GTC ATC 2451 TIT AAA ATG AAG TGG CAA TAC AAC CGA AGG GCT TTC ACA ATG CTA GGT GTG GTC TGG CTG GTG GTC ATC 550 TTC ACA ATG CTA GGT GTG GTC TGG CTG GTG GTC ATC 550 TTC ACA ATG CTA GGT GTG GTG TGG GTG GTG GTG ATC 550 TTC ACA ATG CTA GGT GTG GTG TGG GTG GTG ATC 550 TTC ACA ATG CTA A	297 GTA GGA TCA CCC ATG TGG CAC GTG CAA CAA CTT GAG ATC AAA TAT GAC TTC CTA TAT GAA AAG GAA CAC ATC TGC 3 526 GTA GGA TCA CCC ATG TGG CAC GTG CAA CAA CTT GAG ATC AAA TAT GAC TTC CTA TAT GAA AAG GAA CAC ATC TGC 6	372 TGC TTA GAA GAG TGG ACC CGT GTG CAC CAG AAG ATC TAC ACC TTC ATC CTT GTC ATC CTC TTC CTC CTG 4601 TGC TTA GAA GAG TGG ACC CT GTG CAC CAG AAG ATC TAC ACC TTC ATC CTT GTC ATC CTC TTC CTC CTG 601	447 CCT CTT ATG GTG ATG CTT ATT CTG TAC AGT AAA ATT GGT TAT GAA CTT TGG ATA AAG AAA AGA GTT GGG GAT GGT 56 CT CT CTT ATG CTT ATG CTT ATG CTT ATG CTT ATG GAT GGT ATA CTT TGG ATA AAG AAA AGA GTT GGG GAT GGT 76	522 TCA GTG CTT CGA ACT ATT CAT GGA AAA GAA ATG TCC AAA ATA GCC AGG AAG AAG AAA CGA GCT GTC ATT ATG ATG 575 TCA GTG CTT CGA ACT ATT CAT GGA AAA GAA ATG TCC AAA ATA GCC AGG AAG AAG AAA AAA CGA GCT GTC ATT ATG ATG 8	597 GTG ACA GTG GTG GCT CTC TTT GCT GTG TGC TGG GCA CCA TTC CAT GTT GTC CAT ATG ATT GAA TAC AGT AAT 626 GTG ACA GTG ACA GTG GCT GTC GTG ACA GTG ACA GTG GCT GTC GTG ACA GTG ACA AGT AAT 636 GCA ACA GTG ACA GTG ACA GTG ACA ACA GTG ACA ACA ACA GTG ACA ACA ACA ACA GTG ACA ACA ACA ACA ACA ACA ACA ACA ACA AC	IT GGA TIT TCC AAC TCC	747 ATC TGT AAT CCC ATT GTC TAT GCA TTT ATG AAT GAA AAC TTCGAAAAAAAAAA	105ן אנה אנה אנה אנה אני אינה עינה בינה בינה בינה בינה בינה בינה בינה ב
	•	•											

1200	1275	1296
HTT GAA GTC AAA TTG	TCT	
AAA	AAT	
GTC	GAG	
GA.		
ATT	CTG	
AAC	GAA	
TTC AGT GAT GGC AAC	CTC TTT AGG TCT GAA CTG GCT	
GAT	AGG	
AGT	TTT	
TIC	CTC	
SC.	GCT	
GAA GCA	CIT	
6 69	CAT CTT	
AGA GAG AAT CCA GTG GAG GAA ACC AAA GGA	CGA	
S S S	AAA	
GAA	CTC	
GAG	AAG	
GTG	AAA AAG	
ပ္ပ	AAG	TAA
AAT	GAG AAG	CAT
GAG	GAG	GGG CAT TAA
AGA	ACA	AGT
CIC	CAG	GAC
ပ္ပ	GAA	TTA
TTI	TGT	CCT
126 TTT	201	276 CCT

,

25	20	75	100	125	150	175	200	225	250	275	300	325	350	375	400	425	431
(z.,	H	æ	ωμ	aα	Δ, Δ,	нн	υυ	нн	ပ ပ	ΣE	zz	တ တ	U	×	н	w	
a	>	M	нн	>>	EE	> >	нн	нн	0 0	ΣΣ	တလ	22	>	4	M	z	
Ø	-7	S	ပပ	E4 E4	> >	A A	# #	មេម	ს ს	нн	× ×	တ တ	ပ	×	>	闰	
æ	Ø	æ	[24 [24	о о	H H	> >	шы	нн	>>	>>	121 121	Eu Eu	>	M	E	4	
H	ы	H	fag (ta)	> >	ს ს	i i	K K	нн	KK	44	нн	ပ ပ	4	cc.	н	ы	
1	×	>	E+ E+	ΣΣ	00	3 3	田田	>>	× ×	m m	ΣΞ	нн	Ø	×	z	臼	
z	4	>	нн	× ×	E E	>>	× ×	H H	× ×	× ×	ΣΣ	нн	codon)	Σ	o	တ	
Ħ	æ	*	L L	ပပ	~ ~	>>	н.н	нн	нн	××	E E	٥٥	_ n	F	Ω	œ	
Q	U	ᅜ	H H	нн	田田	ပ ပ	E	E1 E1	2 2	M M	>>	>>	N STO	н	တ	íz,	
æ	а	>	۵ ۵	Eu Eu	> >	14 14		E+ E+	니다	m m	> >	нн	K (No	ပ	Ē4	H	
ы	ы	ч	တတ	44	4 4	ΣΞ	H H	E E	四四	44	==	44	K KKKK	တ	æ	Ø	
н	函	Æ	111	ს ს	нн	E E	K K	× ×	ж ж	нн	E E	E1 E1		z	ia	ы	
œ	Д	z	4 4	ں ن	00	En En	нн	нн	ს ს	* *	a a	нн	2 2	υ	₍	Ħ	
w	EH	U	H H	a a	H H	4 4	ы ы	××	нн	တတ	Æ Æ	ΣE	ыы	Ħ	×	œ	
[24	×	(Eq	တ္တ	3 2	ΣΣ	KK	L L	aa	× ×	ΣΣ	3 3	K K	22	æ	H	×	
QI	>	ы	ပပ	22	E E	ac ac	aa	нн	တ္ထ	四四	00	нн	ΣE	ø	ы	ы	
ы	ч	æ	нн	۵۵	44	zz	aa	>>	ж ж	K K	> >	E E	Eu Eu	æ	Þ	×	
Д	Δι	ы	[14 E4	တ တ	нн	E1 E1	>>	D. D.	H H	טט	44	>>	44	Δι	> .	×	
E	œ	æ	нн	нн	шш	× ×	# #	တ	нн	==	E . E	ם ם	× ×	တ	Δ,	×	*
н	ı	ſΞij	22	zz	E+ E+	aa	23	E E	고고	нн	ны	۵۵	> >	[E4	z	<u>ы</u>	×
Z	œ	н	£ E	امما	> >	32 32	ΣE	38 38	ΣΣ	F F	AA	> >	нн	e	œ	M	U
H	×	ч	> >	H H	> >	××	D1 D4	回田	> >	KK	> >	េខ ខេ	ام به	×	æ	EH	တ
æ	1	>	£ £	E E	A A	ΣΞ	တတ	េធ	ΣΣ	H H	> >	M M	zz	z	H	œ	Ω
a	Ø	ტ	<u>س</u> ا	E+ E+	E E	M M	ប ប	a 되	a a	>>	H H	២២	ပပ	>	တ	ഥ	ы
Σ	н	H	Σ	> >	လ လ	E	> >	ပပ	a a	ارس	>>	Eu Eu	нн	н	[E4	υ	Д
-	56	51	16 1	101	126	151	176	201	226 P	251 8	276 V	301	326	351	376	401	426 P

Alignment of virtual 1036 nucleotide clone with SEQ ID NO 1 and SEQ ID NO 2

ď
77
AGC
O.
Æ
73
~
Æ,
O
m
\sim
Ŷ
U
7
77
Ÿ
U
۲'n
×
Ų
O
Ō
₹.
-
v
Ö
×
v
GGTGGCCCCG
Ė
\sim
Ų
O
ľħ
×
Ÿ
н
۲'n
ж
v
- 1
•
•
S
7
Ą
じり
562
٠,

825	900	975	1036 1050	1125	1200	1275	1296
ATG	AAT	100 T	TGC	AAG	TTG	TCT	
ATG	AGT	AAC	TAT	GCA AAG	AAA	AAT	
ATT	TAC AGT TAC AGT	10 10 10 10	TGT		GTC	GAG AAT	
GTC	GAA GAA	TTT	TT	AAG	GAA	GCT	
SCT SCT	ATT	GGA	ZLZ2 GCA	990	ATT	CTG	
GS SS	ATG	ATT	3' LZLZ21 AA TG TCT GCA GTT	ATG	AAC	GAA	
AAA AAA	ATG	ATT		ATG	ည	TCT	
AAG	CAT	CAA	AAAA	ACA	GAT	AGG	
AAG	GTC	GTG	AAA	ATT	AGT	TTT	
AGG	GTT	ATC	PAAA	GGA	TTC	GCT CTC TTT AGG TCT GAA CTG GCT	
၁၁ ၁၁ ၁၁	CAT	GCT GCT	GAAA	TCA	SC.A	GCT	
ATA	TTC	TIT	TTC	AAT	GAA	CLT	
gga aaa gaa atg tcc aaa ata gcc agg aag aag aaa cga gct gtc att atg atg gga aaa gaa atg tcc aaa ata gcc agg aag aag aaa cga gct gtc att atg atg	GCT CTC TIT GCT GTG TGC TGG GCA CCA TTC CAT GTT GTC CAT ATG ATG ATT GAA TAC AGT AAT GCT CTC TIT GCT GTG TGC TGG GCA CCA TTC CAT GTT GTC CAT ATG ATG ATT GAA TAC AGT AAT	TAT GAT GAT GTC ACA ATC AAG ATG ATT TTT GCT ATC GTG CAA ATT ATT GGA TTT TCC AAC TCC TAT GAI GAI GTC ACA ATC AAG ATG ATT ITT GCT ATC GTG CAA ATT ATT GGA ITT TCC AAC TCC	ATT GTC TAT GCA TTT ATG AAT GAA AAC TTC GAAAAAAAAAA	CAA AGG CAT GGA AAT TCA GGA ATT ACA ATG ATG CGG AAG AAA	GAG AAT CCA GTG GAG GAA ACC AAA GGA GAA GCA TTC AGT GAT GGC AAC ATT GAA GTC AAA TTG	CAT	
1 1 1 1 1 1 1 1	4 55	ATG	E SE	CAT	AAA	CGA	
ATG	1GG	AAG	AAT	AGG	ACC	AAA	
GAA GAA	55 55	ATC	ATG	CAA	GAA	CIC	
E E	512 512 512	ACA ACA	FFF	GC.	GAG	AAG	
GGA AAA GGA AAA	GCT GTG GCT GTG	GTC	ర్ట్ క్ర	cca gca	GTG	AAA	
CAT	TTT	GAT	ATT GTC TAT ATT GTC TAT	ACC TTC TCT	CCA	AAG AAA AAG CTC AAA CGA CAT	TAA
ACT ATT CAT ACT ATT CAT	GCT CTC TTT GCT CTC TTT	GAT	5 E	TTC	AAT	GAG GAG	GGG CAT
ACT	GCT	TAT	ATT	ACC	GAG	GAG	999
69 89 89	GTG GTG	GAA	8 8	AAA	AGA	ACA	AGT
5 5	GTG	AAG	AAT	AAT	CTC	CAG	GAC
TCA GTG CTT CGA TCA GTG CTT CGA	GTG ACA GTG GTG GTG ACA GTG GTG	TTT GAA AAG GAA TTT GAA AAG GAA	ATC TGT AAT CCC ATC TGT AAT CCC	GTA	TCC	TGT GAA CAG ACA	TTA
TCA GIG CIT CGA	GTG ACA GTG GTG 826 GTG ACA GTG GTG	TIT GAA AAG GAA 901 III GAA AAG GAA	ATC TGT AAT CCC 976 ATC TGT AAT CCC	1051 ATA GTA AAT AAA	1126 TTT TCC CTC AGA	TGT	1276 CCT TTA GAC AGT
751	826	901	976	051	126	1201	276

25	20	75	100	125	150	175	200	225	250	275	300	325	350 345	375	400	425	431
							•	.,	• • •	,,		.,	.,.,		•	•	•
Eu (Eu	HH	AA	a a	αa	<u>а</u> а	нн	ပပ	чн	ပ ပ	ΣΣ	22	တ တ	ပ	×	႕	တ	
00	> >	××	нн	> >	H H	>>	нн	라리	0 0	ΣΞ	တ တ	zz	>	æ	×	z	
ω ω	h h	တတ	ပပ	En En	> >	44	E E	E1 E1	თ თ	нн	74 74	တ တ	ပ	×	>	ω	
K K	44	8 8	[24 E4	D4 D4	급 리	> >	ыы	卢႕	>>	>>	គេគ	նո նո	>	×	EI .	Æ	
E E	ㅋㅋ	₽ ₽	[E1 [E1	> >	ပ ပ	ㅂㅂ	× ×	+ +	2 2	44	нн	ს ს	⋖	pc,	H	ы	
니다니	K K	>>	₽ ₽	ΣΞ	αa	3 3	шш	>>	X X	8 8	ΣΣ	нн	တ	Σ	z	ы	
2 2	4 4	>>	нн	× ×	==	>>	> >	卢ᆈ	* *	× ×	ΣE	нн	codon)	×	_U	ဟ	
H H	~ ~	× ×	그 그	ပပ	~ ~	> >	H리	нн	нн	* *	шш	αα	V OP CC	£+	Ω	ĸ	
Ω Ω	ს ს	स्य स्य	нн	нн	ωω	ს ს	£4 £4	E4 E4	3 3	× ×	>>	>>	N ST	н	တ	(Sa)	
~~	<u>а</u>	> >	0 0	E E	> >	리리	00	₽ ₽	17 17	82 82	> >	нн	K (No	ტ	ţ e	ı	
ㅁㅁ	11	11 11	တ တ	44	4 4	ΣΣ	>> >-	E+ E+	шш	44	ππ	4 4	K	တ	A	Ø	
ㅋㅋ	ច្ច	4 4	H H	ပ ပ	нн	H H	× ×	> >-	× ×	+ +	E4 E4	E4 E4	<u> </u>	z	函	ы	
84 8	D1 D2	zz	44	ပ ပ	ပပ	E1 (E1	нн	+ +	ပ ပ	× ×	<u>а</u> а	нн	zz	O	O	Ħ	
တ တ	H H	ပ ပ	LI LI	교괴	6 6	AA	ខាច	× ×	нн	တ တ	4 4	ΣΣ	四日	Ħ	×	ĸ	
E1 E1	× ×	E4 E4	တ တ	33	ΣΣ	8 8	ㅂㅂ	aa	××	ΣΣ	3 3	× ×	zz	ĸ	E	×	
aa	>>	111	ပပ	2 2	E1 E1	K K	aa	шш	တတ	EL EL	ပပ	нн	ΣΣ	œ	函	ы	
ы ы	14 14	4 4	нн		1111	zz	aa	>>	× ×	* *	>>	F F	E4 E4	æ	ഥ	×	
DL DL	а а	H H	£4 £4	တတ	нн	F1 E1	>>	р, р,	H H	ပ ပ	A A	>>	44	щ	>	×	
E4 E4	8 8	44	нн	нн	四四	× ×	= =	တ တ	нн	==	E	00	>+ >+	w	Ω ₄	×	*
нн	14 14	E4 E4	22	zz	H H	aa	33	E4 E4	급교	нн	H H	0 0	>>	Ē	z	ഥ	æ
zz	K K	нн	E+ E+	aa	>>	3 3	ΣΣ	33	ΣE	E+ E+	A A	* *	нн	H	闰	ഥ	Ŋ
H H	× ×	H H	>>	14 14	>>	K K	A 4	면면	> >	2 2	>>	шы	6 1 61	×	œ	E	တ
4 4	нн	>>	E+ E+	ΣΣ	4 4	ΣΣ	တ တ	шш	ΣΞ	H H	> >	K K	22	z	ы	a	Q
aa	44	ပ ပ	44	E E	E E	K K	ს ს	ㅋㅋ	H H	>>	H E	ធធ	ပပ	>	တ	闰	н
ΣE	нн	E E	76 M	<u> > > </u>	တ တ	E1 E1	<u>> ></u>	υυ	<u>a, a,</u>	o o	<u>> ></u>	E4 E4	нн	H.	<u>[14</u>	ن	Д.
Н	26	51	76	101 V	126	151	176	201	226	251	276	301	326	351	376	401	426

Names of IMAGE/Unigene clones

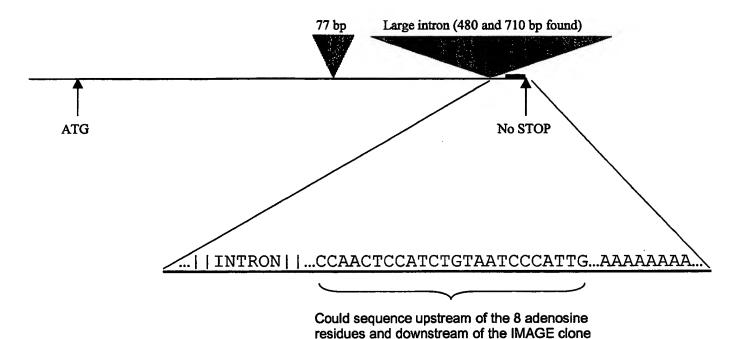
SUBJECT _	m	rje a	Ou	es motelin	4	6A1		
In In	2055518	2	EST 1	9 400 2 3 7 400 2 759 408 7861 Homology Search		source		1 2 3 4 5 6 7
		GB:g4187861				GB:g4001 GB:g4001 GB:g4002	837	
Hit ID GB:g40018 Analysis Results: 28499 GB:g4002759 (D	37 (Detail) Feature Start source 1	0 Stop Desc 432	ent. 100 ription	Description tb27a03.x1 NCI_CGAP_on ://lifetools-e.wh.bayer.com/cgi-bin/	incyte/Life	eTools/2.0/LT_v	/iewer?id=	
sour	etail) 6.2e-43 ure Start Stop ce 3 127 mage/Unigene Ins IGR Indices Les	Descriptions sert Read ngth Length	te670	e05.x1 Soares_NFL_T_G Source (Age/Sex)	-	Homo sapi	ens	20 27
A1308124 I	m+2055532 s.181638 m:2055185 s.181638	14 432 08 432	419	NCI_CGAP_Kid12 libra made from kidney (2 pooled tumors (clear type)) NCI_CGAP_Kid12 libra made from kidney (2 pooled tumors (clear type))	cell	2 (1	2	28 29 30 31 32 33
SIGNED BY	m:2091776 s.181638 AND UNDERSTOOD	404	364	Soares_NFL_T_GBC_S1 library made from po	E	+31	7 16 EYNOLDS /	MILES, MI / SEPT. 95

Location of in-frame STOP codon in sequence upstream of START methionine

SUBJECT NPY-Like Clove and prim	ers		_
	*		1
13			2
### DNA StriderTM 1.2 ###	1/2		3
LGZ1.seq -> Genes			4
DNA sequence 1902 b.p. TGGCCCTCGAGG CTGTGTTCCTTC linear			5
1 TGGCCCTCGAGGCCAAGAATTCGGCACGAGGGGGGGGGG	Met Oher-		6
101 CCTTAGGGAAGAGCAAGGGAAGAACTTTATTTGAACCGCGAACATTTTTTGGTCACTGAGATCGAGTCTCCCAGTGCTTTGGCTTCCCGCCTCTTTATCG 200 M-FRAME 3TOP (FAIL) 201 TGGGTTTGATCCCTGAGCTGCTCTCCTTTCCCGAACCTGCGCGCGACCTAGAGCCCTCCCGCGCGGCTGACTCCAGAGTAGAGGAAGTAGAGCAGCCGAGCGAG	Bloom juice	+ 1/2/105)	
301 CTCCGGCTGGTCCCCCGAAGCCCTCGCTGCCCGCAG ATG CGC AGC CAG TAG CGGGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCGCGGTGGCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCCGCGGTGGCCCCCGCGGTGGCCCCGCGGTGGCCCCGCGGTGGCCCCCGCGGTGGCCCCCGCGGTGGCCCCCGCGGTGGCCCCCGCGGTGGCCCCCGCGGTGGCCCCCGGGGGG			8
393 A ANG COLOGO CIT AAC ATT ACC CCG GAG CAG TTC TCT CGG CTG CCG GAC CAC AAC CTG ACG CGG GAG CAG 465			9
466 TTC ATC GCT CTG TAC CGG CTG CGA CCG CTC GTC TAC ACC CCA GAG CTG CCG GGA CGC GCC AAG CTG GCC CTC GTG 540			10
25 F I A L Y R L R POVININAL Y P E L P G R A FK L A L V 49 541 CTC ACC GCC GTC GTC ATC THC GCC CTC GCC GCC AAT GCT CTC GTG GTG ACC GCC AGC AAG 615	1.1.	AA 66	11 /
50 L T G V L I F A L A L G N A L V F Y V V T R S K 74	C & 7	7 (6	12 ·
616 GCC ATG GTC ACC GTC ACC AAC ATC TTT ATC TGC, TG GTC AGT GAC CTG CTC ATC ACC TTC TGC ATT 690 75 A M R T V T N I F I C 691 CCC GTC ACC ACC ATC CTC CAG AAC ATT TCC GAC AAC TGG GGG GGT GCT TTC ATT TGC AAG ATG GTG CCA TTT GTC 765 100 P V T M L Q N I S D N W L G G A F I C K M V P F V 124		ACC	14
100 P V T M L Q N I S D N W L G G A F I C K M V P F V 124	3		15
- 766 CAG TCT ACC GCT GTT GTG ACA GAA ATC CTC ACT ATG ACC TGC ATT GCT GTG GAA AGG CAC CAG GGA CTT GTG CAT 840 125 Q S T A V V T E I L T M T C I A V E R H Q G L V H 149	CC	TAGA	
841 CCT TTT AAA ATG AAG TGG CAA TAC ACC AAC CGA AGG GCT TTC ACA ATG CTA GGT GTG GTC TGG CTA GTC 915 150 P F K M K W Q Y T N R R A F T M L G V V W L V A V 174	667	FACC	17
916 ATC GTA GGA TCA CCC ATG TGG CAC GTG CAA CAT CTT GAG ATC AAA TAT GAC TTC CTA TAT GAA AAG GAA CAC ATC 990			18
991 TGC TGC TTA GAA GAG TGG ACC AGC CCT GTG CAC CAG AAG ATC TAC ACC ACC TTC ATC CTT GTC ATC CTC TTC CTC 1065	7-	A	19
200 C C L E E W T S P V H Q K I Y T T F I L V I L F L 224 1056 CTG CCT CTT ATG GTG ATG CTT ATT CTG TAC AGT AAA ATT GGT TAT GAA CTT TGG ATA AAG AAA AGA GTT GGG GAT 1140			20
225 L P L M V M L I L Y S K I G Y E L W I K K R V G D 249			21
1141 GGT TCA GTG CTT CGA ACT ATT CAT GGA AAA GAA ATG TCC AAA ATA GCC AGG AAG AAG AAA CGA GCT GTC ATT ATG 1215 250 G S V L R T I H G K E M S K I A R K K R A V I M 274			22
1216 ATG GTG ACA GTG GTG GCT CTC TTT GCT GTG TGC TGG GCA CCA TTC CAT GTT GTC CAT ATG ATG ATT GAA TAC AGT 1290 275 M V T V V A L F A V C W A P F H V V H M M I E Y S 299			23
1291 AAT TIT GAA AAG GAA TAT GAT GAT GTC ACA ATC AAG ATG ATT TIT GCT ATC GTG CAA ATT ATT GGA TIT TOCKAG 1365 300 N F E K E Y D 80 VAS T I K M I F A I V 70 I I G F S N 324		,	24
136 120 APP TO THE GCA THE ATG AAT GAA AAC TEC AAA AAA AAT GET TEG TOT GCA GET TET TAT 1440 325 8 1 C N P 1 V A P M N E N F K K N V L S A V C Y 349			25
1441 TGC ATA GTA AAT AAA ACC TTC TCT CCA GCA CAA AGG CAT GGA AAT TCA GGA ATT ACA ATG ATG CGG AAG AAA GCA 1515		•	26
350 C I V N K T F S P A Q R H G N S G I T M M R K K A 374 "1516 AAG TIT TCC CTC AGA GAG AAT CCA GTG GAG GAA ACC AAA GGA GAA GCA TTC AGT GAT GAC AAC ATT GAA GTC AAA 1590			27
375 K F S L R E N P V E E T K G E A F S D G N I E V K 399 1591 TTG TGT GAA CAG ACA GAG GAG AAG AAA AAG CTC AAA CGA CAT CTT GCT CTC TTT AGG TCT GAA CTG GCT GAG AAT 1665	2.		28
400 L C E Q T E E K K K L K R H L A L F R S E L A E N 424	11-1		29
1666 TCT CCT TTA GAC AGT GGG CAT TAA TTATAACAATATCTTCATAATTA ATG CCC TTC AGA TTG TAA CCCAAAGACAAAATTATTTT 1750 425 S P L D S G H * M P F R L 6		 	30
1751 GAGCAAAGGTCAAATACTCTTTTATTCTTAAG ATG ATG ACA AGA AGA AAA CAA ATC ATG TIT CCA TTA AAA AAT GAC ACG 1831 M M T R R K Q I M F P L K N D T 16		 	31
- 1832 AGG CTA GIC CAA GTG CAG TGA TGATTAAACAATCATTTAACACATTTCTGTGTTCCTTC 1902		 -	32 33
B 2 3		 	34
	".DS & RE		S, MI ' SEPT. 95

Cloning the 3'-end based on IMAGE sequence clues and genomic structure of our novel GPCR

Diagram of IMAGE clone structure and BLAST search strategy



intron be used in a BLAST search to identify the missing 3'-terminal region?

Preliminary genomic structure of novel GCPR



Final electronically assembled novel GPCR (named LGZ1)

	SUBJECT	NRY-L	ike Clove	and pr	iner	<u>}</u>	
		The state of the s		13	-		1
###	DNA Strider TM 1.2 ###						2
 LG	Z1.seg -> Genes			井10,	4, 15, 16	`!	3
		CCCTCGAGG CTGTGTTCCTTC	C linear				4
	1 TGGCCCTCGAGGCCAAGAATTCGGCAC	GAGGAGGCGGGGGAGCCAGAGGCGCCL	AGGACCCTCGCGTGGCGCTCCAGC	ACCCCAGACCGTGGCGGGGCTTCC	100.		5
	101 CCTTAGGGAAGAGCAAGGGAAGAACTT	TATTTGAACCGCGAACATTTTTTGG		CITTGCCTTCCCGCCTCTTTATCC	ı	oby .	6
-	201 TGGGTTTGATCCCTGAGCTGCTCTCCT	TTCCCGAACCT SEES TTCCAGCCT	PAGAGCCCTCCCGCGCGCTGACTY	CAGAGTAGAGGAAGGAAGGAGGA	300	7	8
٠	301 CTCCGGCTGGTCCCCGGAAGCCCTCGC	M R M A	s Q *	->	392		9
•	39 LANG PAGE CIT AAC ATT A	CC CCG GAG CAG TTC TCT CGC P E Q F S R	G CTG CTG CGG GAC CAC AAC L L R D H N	CTG ACG CGG GAG CAG L T R E Q	465 24		10
	466 TTC ATC GCT CTG TAC CGG CTG 25 P I A L Y R L	CGA CCG CTC GTC TAC ACC C	CCA GAG CTG CCG GGA CGC (SCC AAG CTG GCC CTC GTG	540 49	CAA	 -
	541 CTC ACC GGC GTG CTC ATC THE	GCC CTG GCG CTC TTT GCC	AT GCT CIG GIG TIC TAE	STG GTG ACC CCC AGC AAG	615	cfer	12
	616 GCC ATG GC ACC GTC ACC AAC 75 A M RIT V T N	ATC TIT ATC TG, T	EG CTC AGT GAC CTG CTC	ATC ACC TTC TTC TGC ATT	. ' · · · · · · · · · · · · · · · · · ·	CTT6	6 13
	691 CCC OFT ACC ATT CTC CAG AAC	ATT TCC GAC AAC TSS CG	GG GGT GCT TTC ATT TGC A	AG ATG GTG CCA TIT GTC	765	6 AAQ	C 14
	551 CCC OTC ACC ATC CTC CAG AAC 100 P V T M L Q N 766 CAG TCT ACC GCT GTT GTG ACA	I S D N W L ^v G	GAFICE	C M · , V P F V	124 3		15
	125 Q S T A V V T	EILTMT	CIAVERI	I Q G L V H	840 149	CIA	6A16
	841 CCT TTT AAA ATG AAG TGG CAA 150 P P K M K W Q	YTNRRAF	TMLGV	WLVAV	915	6 AA	PC 17
	916 ATC GTA GGA TCA CCC ATG TGG 175 I V G S P M W	CAC GTG CAA CAA CTT GAG A	TC AAA TAT GAC TTC CTA 1 K Y D F L Y	TAT GAA AAG GAA CAC ATC	990 199		18
	991 TGC TGC TTA GAA GAG TGG ACC 200 C C L E E W T	AGC CCT GTG CAC CAG AAG AS P V H Q K I	TC TAC ACC ACC TTC ATC C	TT GIC ATC CTC TIC CTC	1065	TA	19
	1066 CTG CCT CTT ATG GTG ATG CTT 225 L P L M V M L	ATT CTG TAC AGT AAA ATT G			1140		20
	1141 GGT TCA GTG CTT CGA ACT ATT	CAT GGA AAA GAA ATG TOO A	AA ATA GCC AGG AAG AAG	א רינא כרייו ביוורי איוויו איווי	1215		21
	250 G S V L R T I	TIT GCT GIG TGC TGG GCA C	ርል ጥጥር ሮልጥ ርጥጥ ርጥር ሮልጥ ል	ייים אינים מודים באום מיונים מיונים	1290		22
	275 M V T V V A L 1291 AAT TIT GAA AAG GAA TAT GAT	FAVCWAP	P H V V H M	MIEYS	299		23
	JOUN FEREY D	D VASIT I K M I	PAI V _{P-1} QI	I G F S N	1365 324		· 24
	136 man we wind wise Am Reic 325 S I C N P I			SAVCY	1440 349	+ =	25
	1441 TGC ATA GTA AAT AAA ACC TTC 350 C I V N K T F	TCT CCA GCA CAA AGG CAT G S P A Q R H G	GA AAT TCA GGA ATT ACA A N S G I T M	TG ATG CGG AAG AAA GCA	1515 374		26 27
:	1516 AAG TIT TOO CIO AGA GAG AAT 375 K P S L R E N	CCA GTG GAG GAA ACC AAA G P V E E T K G	GA GAA GCA TTC AGT GAT G E A F S D G	GC AAC ATT GAA GTC AAA N I E V K	1590 399		1 28
	1591 TTG TGT GAA CAG ACA GAG GAG 400 L C E Q T E E	AAG AAA AAG CTC AAA CGA C	AT CTT GCT CTC TTT AGG T	CT GAA CTG GCT GAG AAT	1665	,	29
- :	1000 ICI CCT TIA GAC AGT GGG CAT	TAA TIATAACAATATCIICATAAT	TA ATG (CCC TTC AGA TTG T	AA CCCAAAGAGAAAATTATTTT	1750		30
	425 S. P. L. D. S. G. H. L751 GAGCAAAGGTCAAATACTCTTTTTATTC	TTAAG ATG ATG ACA AGA AGA	AAA CAA ATC/ATG TIT CCA	TTA AAA AAT GAC ACG	1831		31
	1832 AGG CTA GTC CAA GTC CAG TGA	M M T R R	K Q I (M F /P	r k n d l	16		32
	17 R L V O V D	T. T. T. T. T. T. T. C. A. C.	ni i immununtti Cigigi i Citi	•	1902		33
	<u> </u>	B 23					34
	CIONED THE PROPERTY OF	1/11	1		אַר יאַ	& REYNOLDS	S / NILES, MI 'SEP
	SIGNED BYWITNESSED AND UNDERS	Thomas are	2/chm	DATE	 . 		
	OBOGG BEEFRENOTS.	POUBL TEX	Letmin	DATE _			· ·

Alignment of final 300 3'nucleotides against SEQ ID NO 1

75	150	225	300	375	450	525	009	675	750	825	006	975	33 1050	108 1125	183 1200	258 1275	300 1296	
TTC	CTC	ည	ည္ဟ	CAG	CCT	ATC	TGC	CTG	GGT	ATG	AAT	JCC J	16C	AAG	TTG	TCT		
CAG 7	GTG (AAG (ATT (GTC (CAT (GTC 1	ATC :	CTC (GAT (ATG 1	AGT 1	AAC 3	TAT	GCA 7	AAA	AAT		
GAG	CTC	AGC	TGC	TTT (GTG	GCA	CAC	TTC	999	ATT	TAC	JCC J	TGT	AAA	GTC	GAG		
990	ည္ဟ	ည္ပ	TIC	CCA	CIT	GTG	GAA	CTC	GTT	GTC	GAA	TTT	GTT	AAG	GAA	GCT		
ACG .	CTG	ACC	TTC	GTG	GGA	CTG	AAG	ATC	AGA	GCT	ATT	GGA	GCA	999	ATT	CTG		
CIG	AAG	GTG	ACC	ATG	CAG	TGG	GAA	GTC	AAA	CGA	ATG	ATT	TCT	ATG	AAC	GAA		
AAC	သည	GTG	ATC	AAG	CAC	GTC	TAT	CTT	AAG	A.	ATG	ATT	TTG	ATG	ည ည ည	TCT		
CAC	၁၅၁	TAC	CTC	JGC	AGG	GTG	CTA	ATC	ATA	AAG	CAT	CAA	GTT	ACA ACA	GAT	AGG		
GAC	GGA	TTC	CTG	ATT	GAA	GGT	TTC	TTC	TGG	AAG	GTC	GTG	AAT	ATT	AGT	TTT		
990	၅၁၁	GTG	GAC	TTC	GTG	CTA	GAC	ACC	CTT	AGG	GTT	ATC	A AA	GGA	TTC	CTC		
CIG	CTG	CTG	AGT	GCT	CCI	ATG	TAT	ACC	GAA	ည္ပ	CAT	GCT	AAA	TCA	60 GC A	GCT GCT		
CTG	GAG	GCT	CTC	GGT	ATT	ACA	AAA	TAC	TAT	ATA	TIC	TTT	TTC	AAT	GAA	CIT		
၁၅၁	CCA	AAT	929	999	TGC	TTC	ATC	ATC	GGT	AAA	K CC	ATT	AAC	85 88	GGA GGA	CAT		
TCT	ACC	ည္ဟ	TTG	CTG	ACC	GCT	GAG	AAG	ATT	TCC	GCA	ATG	GAA	CAT	AAA AAA	CGA	<u>Taa</u> itataacaatatcttcataat <u>Taa</u>	
TTC	TAC	TTT	TCC	TGG	ATG	AGG	CIT	CAG	AAA	ATG	TGG	AAG	AAT	AGG AGG	AC CC	A A A	TTC	
CAG	GTC	CIC	TGC	AAC	ACT	CGA	CAA	CAC	AGT	GAA	TGC	ATC	ATG	5 5	GAA	5 5	TATC	
GAG	CIC	929	ATC	GAC	CTC	AAC	CAS	GTG	TAC	AAA	GTG	ACA	TTT	ઈ ઇ	GAG	AAG	ACA	
၁၁၁	೮೦೦	CTG	TTT	TCC	ATC	ACC	GTG	CCT	CTG	GGA	GCT	GTC	GCA	5 5	GTG	A A	TAT	
ACC	CGA	ည္ဟ	ATC	ATT	GAA	TAC	CAC	AGC	ATT	CAT	TIL	GAT	TAT	ដូ ដូ	5 5 5 5	AAG AAG		
ATT	CTG	TTC	AAC	AAC	ACA	CAA	TGG	ACC	CTT	ATT	CTC	GAT	GTC	TTC	AAT	GAG	S ST	
AAC	၁၅၁	ATC	ACC	CAG	GTG	TGG	ATG	TGG	ATG	ACT	GCT	TAT	ATT	ACC ACC	GAG	GAG	999	
CIT	TAC	CTC	GTC	CTC	GTT	AAG	ည္သ	GAG	GTG	CGA	GTG	GAA	ည	AAA	AGA	ACA ACA	AGT	
වුර	CTG	GTG	ACC	ACC ATG	GCT	ATG	GGA TCA	GAA	ATG	CTT	ACA GTG	AAG	TGT AAT	AAT	5 5	CAG	GAC	
CAG	GCT	ပ္ပ	ပ္ပ		ACC	AAA	GGA	TTA	CTT	GTG	ACA	GAA		GTA	ည္သည္မ	GAA GAA	TTA	
ATG	ATC	ACC	ATG	GTC	TCT	TTT	GTA	TGC	CCT	TCA	GTG	TTT	1 976 ATC	ATA	TII	TGT	5 5	
ᆏ	16	151	226	301	376	451	526	601	919	751	826	901	1 976	34	$\frac{109}{1126}$	184	259 1276	

Final cloned full-length novel GPCR

	BAYER CORPORATION
SUBJECT	Seavencing results of LGZI protish
	[Called It (Taylor Point to cloud P.6 recentor)
1 -	
2	and when !
3	
4	Clare \$1 500 is 8k with one nolymerse
5	1 1) Come 14 1 50 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
6	
7	### DNA Strider™ 1.2 ###
8	·
9	lgz1/pT7Blue#1 sop -> Graphic Map The sequence 1597 by carectrocate contenttaca linear
10	DNA sequence 1587 b.p. caagettgeatg cegtegtttaca linear
11	·
12	
13	165 Eam1105 I 1575 BoeF I 1573 Gd II
14	- 1573 Eae I
15	- 20 Sal 327 Sty I 1559 Ecl 36 I 1559 Ban II
16	13 Sse8337! 291 Hae I 1553 Acces I 1553 Acces I
17	8 NspC1 220 SgrA I 705 Bgl II 1544 BamH I 1521 Bsg I 1521 Bsg I
18	8 Nsp 164 BsiE 521 Dra 768 Bsp1407 1065 Ppu10 1455 Afi 2 Hind III 200 Bgi 450 NspB 695 ApaL 1065 Nsi 1395 Eco57 19
19	T+
20	70 KTB 1921/pT7Blue#1 sop 1587 base pairs Unique Sites 1366 STOP
21	10 11 8
22	a Close # 7 4. Brain a good (same orientation)
23	
24	with the stee (Phe) 917 AA
25	
26	3 clone of 6 11. Mart & good
27	with: atc > Att (16e) 216 ATA
28	I cloud cirry VA-WH
29	
30	in peonta 3.1V flis.
31	
32	162 (Sam 3.1/-)
33	63 3 6 253
34	nink and
	D BY DATE
SIGNED	
	SSED AND UNDERSTOOD BY

BAYER CORPORATION

s	SUBJECT	Ful	leu	yth c	ONA	of Le	721	(NP	y-lik	(L).	,
1				n p	DNA		The				
			S	#1 equenche	sop/pcD1 er™ "Un:	-		t "			
	HindIII (6) KpnI (16) SacI (22) BamHI (24 SpeI (30) BstXI (43) SmaI (6)	3)					PstI (1	113) _[E	CcoRI (13	NotI XmaII XhoI XbaI DraII	(1563) (1573) (1578) I (1578 (1584) (1590) I (1596) (1600)
I nal	Cutters: Apal, Non-Cutters:	Bsp106 & 3	SalI				i, Sacii, Sm	al, Spel, 2	Kbal, Xhol	& XmaIII	
		9	tart	#1sc	op/pcDA3.lTopo "" "Untitled Projec						
of ICZI fr ZIZI #1	tee egg gag ege ace tee egg gag ege ace tee egg gag ege ace Ser Arg Glu Arg The	gca atg cag g	W Ctt and att a	oc eeg gag eag tt	e tet egg etg etg e tet egg etg etg	g cgg gac cac as	ac ccy acy cyg	gag cag tt	c atc gct ctg	tac cgg ctg cga	ccg
of LGZ1 fr 7.21 \$118	ctc gtc tac acc cca	gag ctg ccg gg	ia ogo god aag et ia ogo god aag et	g gee ete gtg ete	e acc ggc gtg ctc	ate the gee et	a ded cractr	gc aat gc	ctg gtg ttc	Tyr Arg Leu Arg tac gtg gtg acc tac gtg gtg acc	Pro ege
of 1621 fr 121	Leu Val Tyr Thr Pro	gag ctg ccg gg Glu Leu Pro Gl	a cgc gcc aag ct y Arg Ala Lys Le	g gee ete gtg ete u Ala Leu Val Leu	acc ggc gtg ctc Thr Gly Val Leu	ate ttc gcc ct Ile Phe Ala Le	g gcg ctc tty u Ala Leu Phe	gc aat gct Ly Asn Ala	ctg gtg ttc	tac gtg gtg acc Tyr Val Val Thr	cgc
#235	age aag gee atg ege age aag gee atg ege Ser Lys Ala MET Arg	acc otc acc aa	C ATC TIT ATC TO	2 bee bee	agt gac ctg ctc	atc acc tto tto	c tgc att ccc	gtc acc atg	ctc cag aac	att too gae aac att too Gac aac	tgg tgg
#352	ctg ggg ggt gct ttc	att tgc aag atg att tgc aag atg att tgc aag atg	gtg cca ttt gte gtg cca ttt gte gtg cca ttt gte	cag tot acc get cag tot acc get cag tot acc get	gtt gtg aca gaa gtt gtg aca gaa gtt gtg aca gaa	atc ctc act atg atc ctc act atg atc ctc act atg	ace tge att ace tge att ace tge att	gct gtg gaa gct gtg gaa gct gtg gaa	agg cac cag g	ga ctt gtg cat o	cct cct
	Leu Gly Gly Ala Phe	att tgc aag atg	gtg cca ttt gtc 'Val Pro Phe Val	cag tot acc got Gln Ser Thr Ala	gtt gtg aca gaa Val Val Thr Glu	atc ctc act atg Ile Leu Thr MET	acc tgc att Thr Cys Ile	gct gtg gaa Ala Val Glu			
WIT	TNESSED AND	UNDERS	TOOD BY	Terre	2/0	In	DAT	E			

BAYER CORPORATION

SUBJECT 1 2 3 PCR of ICEL fc... ttt asa atg asg tgg cas acc aac ogs agg get tic aca atg ets ggt gig gig ggg gig gig ggs gic atc gis gga tes eec atg igg eac gig eas as eat it gag atc as ttT asa atg asg tgg cas tac acc asc cgs agg get ttc acs atg cts ggt gtg gtc tgg ctg gtg gcs GTc atc gTa Ggs tcs Ccc atg tgg Cac gtg cas cas ctt gaG atc ass 02_J*Z*IZI 03_12122 ttt aaa atg aag tgg caa tac acc aac cga agg get tte aca atg cta ggt gtg gte tgg ctg gtg gca gte ate gta gga tea eec atg tgg cac gtg caa eaa ett gag ate aaa #469 ttt aaa atg aag tgg caa tac acc aac cga agg get tte aca atg eta ggt gtg gte tgg etg gtg gca gte ate gta gga tea eec atg tgg eac egt gea eac ett gag ate aaa Phe Lys MET Lys Trp Gln Tyr Thr Asn Arg Arg Ala Phe Thr MET Leu Gly Val Val Trp Leu Val Ala Val Ile Val Gly Ser Pro MET Trp His Val Gln Gln Leu Glu Ile Lys RCR of LGZI fr.. tat gac ttc cta tat gas aag gas cac atc tgc tgc tta gas gag tgg acc agc cct gtg cac cag aag atc tac acc acc ttc atc ctt gtc atc ctc ttc ctc ctt ctc ctc ctt atg Tat gac ttc cta tat Gaa aag gaa cac atc tgC tgc tt 02 1.71.71 യുമമ tat gad the ota tat gam mag gam can ato tgo tgo tta gam gag tgg aco ago cot gtg can dag ato tad aco aco tto ato ott gto ato ote tto oto otg cot ott atg 04_12723 c cet gtg cac cag aag atC tac acc acc tte atc ett gte atc Cte tte ete etg eet ett atg #586 Tyr Asp Phe Leu Tyr Glu Lys Glu His Ile Cys Cys Leu Glu Glu Trp Thr Ser Pro Val His Gln Lys Ile Tyr Thr Thr Phe Ile Leu Val Ile Leu Phe Leu Pro Leu MET RDR of LGZA fr.. gtg atg ctt att ctg tac agt amm att ggt tat gam ctt tgg atm amm gam agm ggt ggg gat ggg tcm ggg att ct cgm act att cat ggm amm gam atg tcc amm atm gcc agg amm gig aig ett att eig tae agt aan att ggt tat gan ett igg ain ang ann ag git ggg gat ggt ten gig ett eGn aet att ent ggn ann gan aig tee ann ata gee agg ang 03 12122 04_12723 gtg atg ett att etg Tae agt aan att ggt tat gan ett tgg atn ang man agn gtt ggg gat ggt ten gtg ett egn act att ent ggn ann gna atg tee ann atn gee agg ang gtg atg ctt att ctg tac agt aaa att ggt tat gaa ctt tgg ata aag aaa aga gtt ggg gat ggt tca gtg ctt cga act att cat gga aaa gaa atg tcc aaa ata gcc agg aag Val MET Leu Ile Leu Tyr Ser Lys Ile Gly Tyr Glu Leu Trp Ile Lys Lys Arg Val Gly Asp Gly Ser Val Leu Arg Thr Ile His Gly Lys Glu MET Ser Lys Ile Ale Arg Lys FCR of LGZI fr... aag aaa cya gct gtc att atg atg gtg aca gtg gtg gct ctc tit gct gtg tgc tgg gca cca tit cat git gtc cat atg atg att gaa tac agt aat tit gaa aag gaa tat gat 03 127.72 aag aaa oga got gto att atg atg gtg aca gtg gtg got oto tit got gtg tgo tgg goa oca tit cat git gto Cat atg atg att gAa tho agt aat tit gaa aag gaa tat gat 04 LZL23 aag aaa ega get gte att atg atg gtg aca gtg gtg get ete tit get gtg tge tgg gea eea tie eat git gte eat atg atg att gaa tae agt aat tit gaa aag gaa tat gat 8820 aag aaa oga got gto att atg atg gtg aca gtg got oto tit got gtg tgo tgg goa oca tit cat git gto cat atg atg att gaa tac agt aat tit gaa aag gaa tat gab Lys Lys Arg Ala Val Ile MET MET Val Thr Val Val Ala Leu Phe Ala Val Cys Trp Ala Pro Phe His Val Val His MET MET Ile Glu Tyr Ser Asn Phe Glu Lys Glu Tyr Asp FOR of LOZI fr. gat git aca atc ang atg att tit get atc gig can att att gga tit tec and tec atc tgt and ecc att git tat gen tat gen tat gan and tit and and and and git tig tet 03 12122 gat gtc Aca atc aaG atg att ttt gct Atc gtg caa att att g 04_LZLZ3 gat gtc aca atc aag atg att tit gct atc gtg caa alt att gga tit tcc aac tcc atc tgt aat ccc att gtc tat gca tit atg aat gaa aac tic aaa aaa git tig tct 05 L2L24 gat gic aca atc and atg att tit get Ale gig can att att gga tit tee ane tee ate igt ant eet et it at gen tit atg ant gan and tit ann and git tig tet #937 gat gto aca ato aag atg att tit got ato gig caa att att gga tit too aac too ato tgi aat ooc att gto tat goa tit aig aat gaa aac tic aaa aaa aat git tig tot Asp Val Thr Ile Lys MET Ile Phe Ala Ile Val Gln Ile Ile Gly Phe Ser Asn Ser Ile Cys Asn Pro Ile Val Tyr Ala Phe MET Asn Glu Asn Phe Lys Asn Val Leu Ser FCR of LGXI fr. gca gtt tgt tat tgc ata gta aat aaa acc ttc tct cca gca caa agg cat gga aat tca gga att aca atg atg ogg aag aaa gca aag ttt tcc ctc aga gag aat cca gtg gag 04 12123 gea git igt tal ige ala gia aat aaa aee tie tet eea gea caa agg eat gga aat iea gga alt aca aig aig egg aag aaa gea aag tit tee ete aga gag aat eea gig gag 05 LZI.24 goa git ig tal igo ala gia aat aaa aco tic ict coa goa caa agg cal gga aat ica gga att aca alg alg ogg aag aaa goa aag tit ico cic aga gag aat oca gig gag #1054 goa git igi tal igo ala gia aal aaa aco lic ici coa goa caa agg cal gga aal ica gga all aca aig aig ogg aag aaa goa aag lil ico cic aga gag aal coa gig gag Ala Val Cys Tyr Cys Ile Val Asn Lys Thr Phe Ser Pro Ala Gln Arg His Gly Asn Ser Gly Ile Thr MET Arg Lys Lys Ala Lys Phe Ser Leu Arg Glu Asn Pro Val Glu FCR of LCEA fr., gas acc sas ggs gas gcs ttc agt gst ggc sac att gas gtc sas ttg tgt gas cag aca gag gag aag aas aag ctc sas cga cat ctt gct ctc ttt agg tct gas ctg gct gag 04 LZIZ3 gaa ace aaa gga gaa gea tte agt gat gge aac att gaa gte aaa ttg tgt gaa eag aca gag gag aag aaa aag ete aaa ega eat ett get ete ttt agg tet gaa etg get ge 05_LZLZ4 gaa acc aaa gga gaa gca tto agt gat ggo aac att gaa gto aaa ttg tgt gaa cag aca gag gag aag aaa aag oto aaa oga cat ott got oto ttt agg tot gaa otg got g £1171 gaa acc aaa gga gaa gca tto agt gat ggo aac att gaa gto aaa ttg tgt gaa cag aca gag gag aag aaa aag cto aaa cga cat ott got oto ttt agg tot gaa otg got gag Glu Thr Lys Cly Glu Ala Phe Ser Asp Gly Asn Ile Glu Val Lys Leu Cys Glu Gln Thr Glu Glu Lys Lys Leu Lys Arg His Leu Ala Leu Phe Arg Ser Glu Leu Ala Glu PCR of LGZI fr. aat to cot tta gac agt ggg cat taa tta taa caa tat ott cat aa tgc cct tda gat tgt aac cca aag aga aaa tta ttt tga gca aag gtc aaa tac tct ttt tat tct taa 04 IZI23 cct tta gac agf ggg cat taa tta taa caa tat ctt cat aa tgc cc 05 LZLZ4 At to cct tta gac agt ggg cat taa tta taa caa tat ctt cat aa igo oot toa gat tyt aac ooa aay aga aaa tta tit iga goa aay yic aaa tac tot tit tat tot \$1288 aat tot oot tta gac agt ggg cat taa tta taa caa tat ott cat aal tge eet tea gat tgt aac eea aag aga aaa tta ttt tga gea aag gte aaa tae tet tit tat tet taa Asn Ser Pro Leu Asp Ser Gly His STP Leu STP Gln Tyr Leu His Asn STP Cys Pro Ser Asp Cys Asn Pro Lys Arg Lys Leu Phe STP Ala Lys Val Lys Tyr Ser Phe Tyr Ser PCR of LOZI fr... taa gat gat gac aag aag aaa aca aat cat gtt too att aaa aaa tga cac gag got agt oca agt goa gtg atg ttt aca aco 05_12124 taa gat gat gac aag aag aaa aca aat cat git too att aaa aaa tga cac gag got agt oca agt goa gig atg itt aca acc aag ggo aat tot goa gat atc cag cac agt ggo В taa gat gat gac aag aag aag aaa aca aat cat git tee att aaa aaa tga cac gag get agt eea agt gea gig atg tit aca aec aag gge aat tet gea gat ate eag eac agt gge STP ASP ASP ASP Lys Lys Thr Asn His Val Ser Ile Lys Lys STP His Glu Ala Ser Pro Ser Ala Val MET Fhe Thr Thr Lys Gly Asn Ser Ala Asp Ile Gln His Ser Gly 05 12124 gge ege teg agt eta gag gge eeg egg tte gaa ggt aag eet ate eet aan eet ete ete ggt ete gat tet aeg egt ace ggt eat eat ean eat tga gtt taa ann ege gge oge tog agt eta gag gge eeg egg tte gaa ggt aag eet ate eet aac eet ete egg ete gat tet aeg egt ace ggt eat eat eac eat eac eat tga gtt taa aec ege Gly Arg Ser Ser Leu Glu Gly Pro Arg Phe Glu Gly Lys Pro 1le Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr Arg Thr Gly His His His His His STP Val STP Thr Arg

CROSS REFERENCES

Alignment of LGZ1 clone DNA sequence with SEQ ID NO 1

93 75 25	150	243 225 75	318 300 100	393 375 125	468 450 150	543 525 175	618 600 200	693 675 225	768 750 250	843 825 275
TTC	D L	ည် ရှင်	200 200	CAG CAG Q	CCT	ATC ATC I	ည်း ၁၅၁	CTG CTG	GGT GGT	ATG ATG M
CAG CAG		AAG K K	ATT ATT I	GTC GTC V	CAT CAT H	512 512 5	ATC ATC I	ភ ភ ក ភ ភ	GAT D	ATG ATG M
GAG GAG	1 25	8 S S S S S S S S S S S S S S S S S S S	TGC	TTT TTT F	GTG GTG V	& GC & C & A	25 25 FE	TTC	999 999 999	ATT
999 🚜		ည္ပည္တို့ ဗုပ္ပ	TTC TTC	CCA CCA P	CTT	org org >	GAA GAA	5 5 ¹	GTT GTT	GTC OTC
ACG T		P CC	TTC TTC F	GTG V	664 664 664	OTO TIG	AAG R R R	ATC	AGA R	GCT A
CTG CTG		GTG GTG V	ACC ACC	ATG ATG M	CAG CAG	7 7 8 7 7 6 6	GAA E	6 6T 5 5 >	AAA AAA	CGA R
AAC N		GTG GTG V	ATC ATC I	AAG AAG K	CAC CAC	5 5 >	TAT TAT Y	i i i	AAG K	AAA K
CAC CAC		TAC	STO TO TO	767 767	AGG AGG	ore ore	CTA	ATC	ATA ATA I	AAG AAG K
GAC D D		TTC	CTG CTG	ATT ATT I	GAA GAA	P P P	TTC	TIT	166 166	AAG K
990 8		GTG GTG V	GAC D	TTC TTC F	GTG V	A TO TO	GAC DD	P S S	r C	AGG R R
oro oro		CTG TG TG	AGT AGT S	GCT A	GCT A	ATG M	TAT TAT Y	A ACC	GAA GAA	DD €
CTG CTG		4 GCT	CTC CTC	66T 66T	ATT	A CA	AAA K	TAC	TAT	ATA ATA I
200 CGG		AAT AAT	929 929	999 999 999	146 166 160	TTC	ATC ATC I	ATC	199 199 199	A AAA
TCT TCT S		ပ္ပ ပ္ပ	TTG TTG L	CTG CTG	ACC T	GCT T GCT	GAG GAG	AAG K	ATT	TCC TCC
TTC		TTT	TCC TCC	TGG TGG	ATG M	A GG	CTT	CAG	AAA K	ATG ATG
CAG CAG		ភ្ជុំ ភ្ជុំ	ညီ ၁၅၁	AAC N	ACT T	8 8 ×	A CAS	CAC	AGT	GAA GAA
GAG GAG		9 9 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	ATC	GAC GAC D	2 2 1 5 5 1	AAC N	CAA C	GTG STG	TAC	AAA K
900 d		7. 7. 1. 1.	TTT TTT E	3 1 1 2 C	ATC	2 2 E	GTG GTG	TO CO	CTG CTG	450 6645 6745
ACC T		ပ္ပ ပ္ပ	ATC	ATT ATT I	GAA GAA	TAC	CAC CAC	AGC S AGC	ATT ATT I	CAT CAT H
ATT ATT I		TTC	AAC N	AAC N	ACA T	e e o	TGG TGG	ACC T	CIT	ATT ATT
AAC N N	800 8	ATC	ACC T	CAG	GTG STG >	17GG ¥	ATG ATG M	TGG TGG	ATG ATG	ACT
CTT CTT L	TAC Y	GGC GTG CTC GGC GTG CTC G V L	GTC V	ATG CTC ATG CTC M L	GTT GTT	AAG K	TCA CCC TCA CCC S P	TTA GAA GAG TTA GAA GAG L E E	GTG OTG	CGA R
	13 2	org org	CGC ACC CGC ACC R T	ATG M	GCT A	ATG ATG M	TCA	GAA	ATG ATG M	CHI
ATG CAG ATG CAG M Q	S G 4	ပ္ပို့ ဗ	8 8 8		TCT ACC TCT ACC	AAA AAA K	669 664	TTA	CTT	GTG OTG \
ATG M	76 ATC 26 I	2 2 2 L	244 ATG 226 ATG 76 M	319 GTC 301 GTC 101 V	TCT TCT S	TTT TTT	GTA GTA	7.07 7.07 7.07	CCT	S TCA
1 19 1	76	151	244 226 76	319 301 101	394 376 126	469 451 151	544 526 176	619 601 201	694 676 226	769 751 251

GTG TGC TGG GCA CCA TTC CAT GTT GC CAT ATG ATT GAA TAC AGT AAT GTG TGC TGG GCA CCA TTC CAT GTT GTC CAT ATG ATT GAA TAC AGT AAT V C W A P F H V V H M M I E Y S N

CTC TTT GCT G

844 GTG ACA GTG GTG GCT 826 GTG ACA GTG GTG GCT 276 V T V V A

993 975 325	1068 1050 350	1143 1125 375	1218 1200 400	1293 1275 425	1335 1296 431
TCC TCC	ည်း ၁၈ ၁၈	AAG AAG K	TTG	TCT TCT S	
	TAT	600 4 608 4	AAA K		
TCC AAC TCC AAC S N	TGT	AAA R	215 215 >	GAG AAT GAG AAT E N	
TTT TTT F	GTT	ACA ATG ATG CGG AAG ACA ATG ATG CGG AAG T M M R K	GAA E		
ATT GGA ATT GGA I G	60 A	ATG CGG AAG ATG CGG AAG M R K	AAC ATT AAC ATT N I	STO 1	
	TCT	ATG ATG	GAT GGC AAC ATT GAT GGC AAC ATT D G N I	AGG TCT GAA CTG GCT AGG TCT GAA CTG GCT R S E L A	
	GTT TIG GTT TIG V L	ACA ATG ACA ATG T M	ပ္တစ္တစ္တ	TCT TCT S	
CAA	GTT	ACA	GAT GAT D	AGG R	
GTG GTG	AAT N	ATT ATT I	AGT S	TIT TIT	
GCT ATC GTG CAA ATT GCT ATC GTG CAA ATT A I V Q I	GCA TTT ATG AAT GAA AAC TTC AAA AAA AAT GCA TTT ATG AAT GAA AAC TTC AAA AAA AAT A F M N E N F K K N	TCA GGA TCA GGA S G	GTG GAG GAA ACC AAA GGA GAA GCA TTC AGT GTG GAG GAA ACC AAA GGA GAA GCA TTC AGT V E E T K G E A F S	GCT CTC TTT GCT CTC TTT A L F	
GCT A	A A A	TCA	60 6 60 4	A GCT	
TTT TTT	111C	GGA AAT GGA AAT G N	GGA GAA GGA GAA G E	CAT CTT CAT CTT H L	
ATT ATT I	A PAC	69 69 49 69	69 S	CAT	A.T.
ATG ATG	GAA EAA	CAT	A A A	R CG M	CATA
GTC ACA ATC AAG ATG ATT TTT GTC ACA ATC AAG ATG ATT TTT V T I K M I F	ATG AAT ATG AAT M N	CAA AGG CAT CAA AGG CAT Q R H	GAA ACC AAA GAA ACC AAA E T K	CTC AAA CGA CTC AAA CGA L K R	TCTT
ATC	ATG ATG		GAA EAA	5 5 1	AATA
GAT GTC ACA GAT GTC ACA D V T	TTT	CCA GCA CCA GCA P A	GTG GAG GTG GAG V E	AAA AAG AAA AAG K K	TAAC
GTC V		50 00 m	STG STG	AAA	TTA
GAT D	GTC TAT GTC TAT V Y	होता इत्य	AAT CCA AAT CCA N P	AAG K	TAA TAA
TAT GAT TAT GAT Y D	5 5 >	TTT	AAT	GAG GAG	CAT
TAT	ATT	F ACC	GAG	GAG GAG	999 9
GAA		AAA K	AGA R	ACA T	AGT
919 TTT GAA AAG GAA 901 TTT GAA AAG GAA 301 F E K E	994 ATC TGT AAT CCC ATT 976 ATC TGT AAT CCC ATT 326 I C N P I	1069 ATA GTA AAT AAA ACC TTC TCT 1051 ATA GTA AAT AAA ACC TTC TCT 351 I V N K T F S	1126 TTT TCC CTC AGA GAG AAT 1126 TTT TCC CTC AGA GAG AAT 376 F S L R E N	1219 TGT GAA CAG ACA GAG GAG AAG AAA AAG CTC AAA CGA CAT 1201 TGT GAA CAG ACA GAG GAG AAG AAA AAG CTC AAA CGA CAT 401 C E Q T E E K K K L K R H	1294 CCT TTA GAC AGT GGG CAT TAA TTATAACAATATCTTCATAAT 1276 CCT TTA GAC AGT GGG CAT TAA 426 P L D S G H *
GAA	TGT	GTA V	25. 25. 25.	GAA GAA	TTA
TTT TTT F	ATC ATC	ATA ATA I	TIT	TGT	TOO T
919 901 301	994 A7 976 A7 326 I	1069 1051 351	1144 TT 1126 TT 376 F	1219 TC 1201 TC 401 C	1294 CC 1276 CC 426 P

Alignment of LGZ1 clone amino acid sequence with SEQ ID NO 2

25	20	75	100	125	150	175	200	225	250	275	300	325	350	375	400	425	431
եւ ե ւ	H H	A A	D1 D2	00	<u>а</u> а	нн	ပ ပ	H H	ს ს	Z Z	Z Z	တ တ	ပ ပ	× ×	111	တ လ	
00	> >	× ×	нн	> >	нн	> >	нн	11 11	0 0	ΣΣ	လ လ	2 2	× ×	4 4	K K	zz	
шш	нн	တ တ	ပ ပ	£1 £1	> >	4 4	шш	चि चि	ပ ပ	нн	× ×	တ တ	υυ	× ×	> >	ыы	
K K	4 4	~ ~	ᄄ	о, о,	H H	>>	回田	нн	>>	> >	田田	E4 E4	>>	× ×	阿田	A A	
H H	Li Li	F F	ᄄᄪ	> >	ပ ပ	႕리	× ×	нн	K K	A A	нн	ს ს	A A	~ ~	нн	卢니	
卢卢	× ×	>>	H H	X X	00	3 3	ыы	> >	× ×	~ ~	ΣΣ	нн	တ လ	ΣE	2 2	ω ы	
2 2	4 4	> >	нн	* *	E E	> >	× ×	12 12	* *	× ×	ΣE	нн	H H	ΣΣ	ပ ပ	တ တ	
шш	K K	>> >>	H H	ပ ပ	& &	> >	LI LI	нн	нн	× ×	# #	a a	>>	E+ E+	۵ ۵	m m	
۵ ۵	ს ს	[E4 [E4	i i	нн	回田	ပ ပ	E4 E4	[24 [24	3 3	× ×	> >	> >	zz		လ	[14 [14	
8 8	a a	> >	0 0	[24 [24	> >	a a	ام م	E+ E+	i i	a a:	> >	нн	X X	ပ ပ	[24 [24	L L	
H H	нн	нн	လ လ	A A	4 4	E E	× ×	£+ £+	田田	4 4	H H	A A	× ×	တ တ	A A	A A	
HH	ы ы	A A	12 12	ပ ပ	нн	₽ ₽	M M	× ×	× ×	нн	Eu Eu	[Eq. [Eq.	E4 E4	2 2	шш	ii ii	
K K	<u>а</u> а	zz	A A	ပ ပ	ပပ	E1 E1	нн	нн	ს ს	× ×	a a	нн	2 2	ပ ပ	ပ ပ	= =	
တ တ	H H	ပ ပ	11 11	12 12	E E	A A	ы ы	× ×	нн	တ တ	4 4	ΣZ	ωы	E E	K K	K K	
[14 E4	× ×	E4 E4	လ လ	3: 2	ΣΣ	K K	H H	αa	× ×	ΣΣ	3 3	× ×	zz	K K	E+ E+	K K	
00	> >	H H	ပ ပ	2 2	F F	K K	00	# #	လ လ	ωω	ပ ပ	нн	ΣΣ	00	шш	11 11	
回回	i i	A A	нн	۵ ۵	다 다	zz	αa	> >	* *	M M	> >	E+ E+	E4 E4	4 4	ыы	* *	
4 4	Δ. A.	H H	E4 E4	တ တ	нн	₽ ₽	>>	οι · οι	14 14	ပ ပ	A A	> >	4 4	а а	>>	* *	
H H	K K	4 4		нн	ыы	× ×	# H	တ တ	нн	æ æ	[14 E1	۵ ۵	× ×	တတ	D4 D4	* *	* *
нн	12 12	[21 [21	zz	2 2	E	aa	3 3	E+ E+	H H	нн	12 12	۵۵	>>	E4 E4	2 2	шы	E E
zz	8 H	нн	EH EH	αa	>>	2 3	ΣΞ	33	ΣΣ	EH EH	A A	× ×	нн	E4 E4	ខាគា	ωы	ပ ပ
11 11	× ×	H H	> >	17 17	>>	* *	a a	ស ស	> >	æ æ	> >	ωи	a a	M M	cc cc	₽ ₽	တ တ
A A	нн	> >	F F	ΣΣ	A A	ΣΞ	လ လ	ыы	ΣΣ	H H	> >	M M	z z	zz	нн	00	D D
00	4 4	ပ ပ	K K	e e	e e	M M	ပ ပ	H H	11 12	> >	H H	ខាត	ပပ	> >	တ တ	ច ម	H H
ΣE	нн	타타	ΣΣ	<u>> ></u>	8 8	151 F	176 V V	00	226 P	တ တ	276 V V	[21 E1	нн	нн	376 F	00	<u>6</u> 6
7	26	51	16	101	126	151	17(201	22(251	27(301	326	351	37(401	426

Alignment of assembled fragments composing LGZ1 with SEQ ID NO 2

}		5GA1						3,D A CT #2	S NACE #2				300 nt 3'-end
55	50	75	100	125	150	175	200	225	250	275	300	325	350
E (24	11 11	AA	[A A	0 Q	D; D.		υυ	H H	ဖ ဖ	ΣΣ	zz	တ တ	υυ
00		× ×	Н Н	> >	# #	> >	нн	14 H	Ω Ω	ΣΣ	တတ	2 2	× ×
<u>면</u> 데	H H	w w	U U	Eu Eu	 > >	A A	H H	E4 E4	ပ ပ	нн	→ →	တ ဟ	ပ ပ
W W	AA	8 8	F E GA2	۵. ۵.	급급	> >	ច ម	교교	> >	> >	回回	E4 E4	> >
H H	교교	E E	F E G	> >	ს ს	그 그	× ×	нн	K K	4 4	нн	ပ ပ	A A
니니	M M	> >	I T F I T E 234 bp fragment –	E E	αa	3 3	ខាគ	> >	× ×	64 64	ΣΣ	нн	တ တ
zz	44	> >	I I d bp fi	× ×	==	> >	× ×	нн	× ×	K K	ΣE	нн	ㅂㅂ
# #	E L P G R E L P G R b fragment – NSAN.seq/GA1	× ×	L L 23	υυ	KK KK	> >	11 11	нн	нн	× ×	==	aa	> >
ο ο	G G SAN. See	E4 E4	дд	нн	ыы	ပ ပ	E	E4 (E4	3 3	× ×	> >	> >	z z
KK KK	4 D D		Ω Ω	E4 E4	> >	卢႕	۵ ۵	E+ E+	нн	ec ec	> >	нн	× ×
111	r r r	1 1 2	လ လ	4 4	A A	ΣΣ	× ×	F F1	កក	4 4	==	4 4	× ×
卢卢	E E F	L F G N A L	PH	<u>ဖ</u> ဖ	нн	E+ E+	××	* *	> >-	нн	[24 [24]	E1 E1	fa fa
~ ~	P P 161	N N	A A	ს ს	ပ ပ	E4 E4	нн	нн	დ დ	X X	D4 D4	нн	2 2
တ ဟ	E+ E+	ပ ပ	22	H H	H H	4 4	កាចា	* *	нн	တ တ	4 4	ΣΣ	ыы
E4 E4	× ×	E1 E1 (တ လ	3 3	ΣΣ	K K	H H	a a	× ×	Σ·Σ	8 8	× ×	zz
αa	>>	다 다	ပ ပ	D N D N 2 (con't)	V T E I L T V T E I L T 234 bp fragment - GA2 (con't)	K K	00	# #	တ တ	阿田	ပ ပ	нн	ΣΣ
ចច	H H	4 4	нн	ο ο 24. 24.2 (α	1 L	zz	a a	> >	* *	K K	> >	E1 E1	E4 E4
ابم بم	D1 D1	H H	E4 E4	N I S N I S 234 bp fragment – GA	H H 1	₽ ₽	>>	D1 D2	нн	ს ს	A A	> >	4 4
₽ ₽	24 24	4 4	нн	I I	Fagme		==	တ လ	HH	жж	<u> </u>	Ω Ω	* *
нн	H H	E E	ZZ	N N A	14 bp)	00	3 3	F F	ㅂㅂ	нн	다다	Ω Ω	> >
2 2	22 CZ	нн	T Τ 141, α	992	> > ~	≥ ≥	ΣΣ	2 3	ΣΣ	EH EH	4 4	* *	нн
卢ᆈ	* *	дд	v v	H H	>>	M M	Q. Q.	阿田	> >	K K	> >	ខាច	O. O.
4 4	교리	> >	T	ΣΣ	A A	ΣΣ	တ တ	E E	ΣΣ	ㅂㅂ	> >	××	2 2
a a	4 4	ს ს	ent -	e e	E E	× ×	ပ ပ	ㅂㅂ	H H	> >	HH	阿阳	ပပ
H	26 I	51 T T	76 M R T V T N M M R T V T N M M R T V T N M M M M M M M M M M M M M M M M M M	101 V	126 \$	151 F	176 V V	201 C	226 P	251 S	276 V V	301 F	326 I
			16.										

300 nt 3'-end 375 400 425 431 M M 1 တ တ 22 4 4 XX XX ωи > > **Æ** Æ X X មាធ **K** K ㅁㅁ нн छ छ ΣΣ 22 ပ ပ လ လ ΣZ ₽ # 0 0 **A A** တလ E4 E4 д д ဖ ဖ [24 [24 တ တ 4 4 44 មេខ H H zz ပ ပ **=** = ს ს HH ×× **84** 84 8 8 E E MM αa 四田 ᆸᆸ **121** 121 **K K** A A ᅀ > > K K တလ ם K K ធា ធា zz **= =** ᄪᄪ 四田 ს ს HH 回日 K K **8 2** E E က က 11 aa **Δ** Δ ZZ 卢괴 > > တလ ωω 426 P 401 C 376 F

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked.
☐ BLACK BORDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
☐ FADED TEXT OR DRAWING
☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

IMAGES ARE BEST AVAILABLE COPY.

OTHER:

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.